

2 April 1981

Personal privacy is absolute privacy

How strictly should the rules of privacy apply to the huge banks of personal data now accumulating in the computers of public and private organizations? The European Science Foundation has raised in the past few weeks an intriguing twist on the conventional view that privacy means nothing unless it is absolute. While acknowledging that private persons are entitled to protection, the foundation says that researchers should have access to banks of otherwise confidential data in the interests of "freedom of research". Thus, in its "statement concerning the protection of privacy and the use of personal data for research", the foundation says that governments planning to model legislation along the lines suggested by the Council of Europe's convention on the protection of privacy should be sure to permit access by researchers to otherwise secret data banks. The issue thus raised is intricate, and potentially explosive. The statement in which it is embodied is a much diluted version of an earlier proposal which the General Assembly of the foundation in 1979 would not endorse. Its origins lie in the deep conviction of many social scientists that, as things are, valuable information of potentially great value in the understanding of society is denied to them. But it is unlikely that governments will respond as they would wish to the foundation's proposals. Moreover, there are good reasons why they should not do so.

Of the potential benefits of access by researchers to confidential data, there is very little doubt. In most industrialized countries, publicly compiled data banks would be gold mines for social scientists. Access to computerized health records would make possible far more accurate assessments of the pattern of disease, morbidity and sheer malingering than are possible at present. Access to census data would permit more sensitive studies of marriage patterns, the incidence of social deprivation and of poverty, than the official census takers usually embark on. Access to income tax records would make possible more accurate estimates than those now available of the incidence of wealth within communities — and perhaps even the incidence of tax evasion. In many respects, these data would be of superior quality to that which can be gathered by social scientists themselves using the more traditional methods of survey research. And there is at least some reason to suppose that the benefits would be substantial. More would be learned about the behaviour of different groups within communities than is possible with existing techniques. The frustration of the social scientists that these data are not available to them is understandable. But it is a frustration with which they will have to live for many years to come.

In the past several years, governments have only slowly woken up to the need for protecting personal data held in computer files. Nobody should underestimate the difficulties, the chief of which is the ease with which large quantities of data of a personal character can be stored and then made accessible. In the old days, ordinary people could rely on the difficulty of extracting meaningful data from shelves full of paper files as an assurance of confidentiality. But even where governments have recognized that their citizens are entitled to know that their private affairs will not be broadcast to all who seek to know, problems have arisen in deciding what limits should be placed on the transfer of information from one public authority (say the income tax department) to another (say the police). In equity, there are the strongest reasons why even government departments should be constrained within more or less watertight compartments. It will be a long time before most citizens in advanced societies can be

assured of that. Indeed, outside the United States, enlightened by the Freedom of Information Act, few governments admit the principle that citizens have a right to check data about themselves that may be recorded in the public files. The Council of Europe's convention on privacy is still very much an ideal. It will be an uphill struggle to ensure that it is widely adopted.

This is only one of the reasons why the European Science Foundation's proposal is, at this stage, unhelpful. Rightly or wrongly, it can only raise in ordinary people's minds the fear that supposed guarantees of privacy are not to be taken seriously. And although the foundation's statement suggests that scientists given access to confidential files should be required to follow rules of good conduct, and that they should not make identifiable personal data public without the consent of all those concerned, assurances to this effect will be unpersuasive and legally unenforceable. For where legislation of privacy is contemplated, the onus of responsibility for keeping confidentiality rests with identifiable individuals, the custodians of the sensitive data banks. It is unthinkable that this responsibility could be effectively derogated to whichever academic researchers chose to interest themselves in particular data banks — and unlikely that the researchers would be willing to accept it. The European Science Foundation's proposal is therefore ill-timed.

Scenarios on energy

Last week was a milestone in the short history of the International Institute for Applied Systems Analysis in Vienna, the research institute to which seventeen national academies now provide financial support. For the institute at last published a summary report of its gigantic energy study, on which more than 140 people have been engaged, off and on, since 1973. Readers of the document, *Energy in a Finite World* (Ballinger Publishing Company, \$16.50; Harper and Row Ltd, £12.50) will be of two kinds — those who want to know whether they (or their grandchildren) will be shivering from cold in the year 2030 (when the period covered by the bulk of this study ends) and those who wish to know — perhaps because they have to help find the funds to keep it going — whether the institute's reputation will be enhanced by the appearance of its first important document. On both counts, perhaps in the nature of the questions, the answers are disappointing, even frustrating. As energy scenarios go, the institute's study is more cheerful than most others, but its optimism is so hedged about with qualifications that many fond grandparents will nevertheless think of converting their legacies into string vests and other ways of keeping warm. And the study itself, however impressive its credentials — the leader of the team was Professor Wolf Häfele, a heavyweight in all senses of the word — is deficient in that it fails to acknowledge that the problem of energy is no longer technical but political.

It is now conventional to say that there has been an energy problem since the second half of 1973, when the members of the Organization of Petroleum Exporting Countries (OPEC) abolished by the stroke of a pen what had been until then the most valuable of the developed world's energy resources — crude oil at a price of \$1.50 a barrel. (The same oil costing \$32.50 a barrel seems a very different commodity to those who use it.) The institute's study, urging that those preoccupied with the energy



crisis should take a long view, avoids the trap of suggesting that the energy problem came of age in 1973. The truth is, of course, that it has been with us always. Was not the case for nuclear power in the 1950s based on the assumption that oil reserves would be exhausted in a couple of decades or so? Did not the great W.S. Jevons, leader of the Manchester Liberal School in the 1860s, bemoan the condition of Britain when the coal ran out? For the past 4,000 years, since the first machines were invented, communities of different kinds have found themselves in difficulties for lack of the means for driving their machines. If there is anything special about 1973, it is that so many complacent governments of oil-importing states were prepared to believe that oil-producing states would go on indefinitely selling their oil in whatever quantities were demanded and at a knockdown price. This perspective is a plus for the institute.

So what is special about the present period? The institute's report deals with the period 1975–2030 and starts with a daring but realistic assumption — population growth in the world as a whole during that period will be faster than in any previous time, but will afterwards begin to level off. The assumption is daring because most comparable studies have assumed that population growth will continue and be what is loosely (and falsely) called "exponential". Many doomsayers will complain that the institute is hopelessly optimistic. But the assumption is realistic for at least two reasons. Demographic history shows that sooner or later all communities learn how to balance births against deaths (and the doubling of world population in the next 50 years is less than the threefold increase in Britain between 1815 and 1865, when that demographic transition was more or less accomplished). The second reason is that those states that fail to keep their populations within bounds will end up in catastrophe. This is another plus for the institute.

So too is what the institute has to say about a number of other features of the business of energy forecasting. Technology will indeed make economic activity less energy-intensive than at present. Energy consumption per unit of Gross Domestic Product will continue to fall in developed countries as prosperity (Gross Domestic Product per head of population) continues to increase. And developing countries will not repeat the excesses of the states now industrialized, with the result that they will be able to attain (in half a century) prosperity not very different (outside Africa and South-East Asia) from that now taken for granted in Western Europe without ever having to consume more than 1 watt of energy for each annual dollar's worth of product. These assumptions obviously take the edge off the calculation of how supply and demand for energy can ever be matched. Once the need to worry about what will happen when developing countries consume energy as intensively as developed countries do now is done away with, the collective pressure on resources is diminished and the problem of supply and demand becomes manageable. This part of the institute's argument is well documented and well argued. Only those with a vested interest in believing that the worst will always happen will protest. So this is very much a plus for the institute.

How then does the future look? That is the question futurologists inevitably ask. Mercifully, the institute answers the question in the right spirit, thereby avoiding the snares which have trapped its predecessors. The flaw in most studies of this kind is that a scenario, while not a prophecy, is a hint of what may happen "other things being equal". The institute's study is more imaginative. The objective of a scenario, its study might have said explicitly, is not to show what the future will be like — for none of us can know — but to show that, on plausible and realistic assumptions, the future is possible. The problem, in other words, is to prove what the mathematicians would call an existence theorem. When that is done, extraneous influences — market forces for one, ingenuity for another — will do their work. And once, it has been demonstrated that getting from A (here) to B (there) is possible, it is best not to try to predict how it will be done. Apart from the safe prediction that for most of the next half-century fossil fuels (oil included) will continue to predominate, and from the guesses that global energy production will increase

dramatically in the next 20 years (by a factor of between 1.6 and 2.0) and fantastically (by a factor of between 2.6 and 4.2) by 2030, the institute gives few hostages to fortune. This is the good sense to be expected of a bunch of self-confessed natural scientists. Yet another plus for the institute, it will be concluded.

But how do we get from A to B? That is not a question in futurology, but a question for the present. What do we do now to make the possible future happen? The institute's report takes several novel and stimulating tacks. It has a cheerful account of the virtues of hydrogen as a convenient alternative to electricity as a secondary fuel (and urges that there should be more deliberate attempts to trap the hydrogen released from the deformation produced in the roots of plants in associations with nitrogen-fixing bacteria). The report is lukewarm but not frigid on biomass, futuristically open-minded about energy-collecting satellites in synchronous orbits and philosophical about nuclear power in both its forms — fission and fusion; it's a matter of time, and of giving people a chance to believe that those technologies can be made safely to work. Hurrah, it will be said; maybe there are several routes from A to B. That, for the institute, is a plus of a kind — perhaps half a plus. For these imaginative (but well written) parts of the summary report do not tell us what to do to get from A to B. Putting money into biomass (another name for wasteful forestry) or hydrogen research is only half an answer. For the truth about energy has become political, and will remain so at least until 2030.

Internationally, the proof that energy is a political matter came in 1969, with the *de facto* revocation by Libya of the licences it had earlier given to the multinational oil companies operating there; OPEC's resolve (in 1973) to put flesh on the bare bones of its long-declared intention to secure a fair share of the rent (in Adam Smith's sense) for its oil followed not long after. Domestically, until this year, the United States has been the worst offender, for political reasons charging too little for gasoline and other petroleum products, in the process giving its citizens the illusion that oil is rather like atmospheric oxygen — breathe in (exert demand) and the vacuum will be filled. President Reagan in the past two months has succeeded where his three immediate predecessors (Carter, Ford and Nixon) failed in deregulating the price of natural gas and crude oil in the United States. The short-term consequences have been dramatic — United States imports of petroleum have fallen to just over 3 million barrels of oil a day — 60 per cent of what they were two years ago. Even so, the prices of oil products in the United States are still less than elsewhere — no American administration, not even Reagan's, would risk a thoroughly free market (freedom for exporters and what that entails). The Americans, however, are only the worst and not the only offenders. The British, having found oil in the North Sea, are busily using the tax revenues arising as a poultice for persistent economic ills. Meanwhile, the moderation of OPEC in the past few months owes as much to the continuing war between Iran and Iraq, and the strengthening military relationship between Saudi Arabia and the United States, as to an explicit decision about the price oil can now command. Nobody is much concerned about Mr Willy Brandt's new-found constituency, the developing countries that happen to have no oil.

The institute's report acknowledges that such extraneous factors must influence the pattern of energy consumption in the decades ahead. Its good cheer is based on the assumption that people will behave rationally, on energy pricing, on the development of nuclear power, on a flexible strategy for research on alternative sources of energy, and in the avoidance of war. The caveats are sensible. It follows, however, that what the world now needs is not an existence theorem but a prescription for the avoidance of irrationality. By its nature, the Vienna institute is ill-equipped to help. Its members are, after all, predominantly natural scientists. *Energy in a Finite World* will bring credit to the institute — and would have brought more if it had seen the light of day some years ago. With the passage of the past eight years, however, the problem of energy has been transformed from one of technical analysis to a set of problems in politics, international and domestic. Who will provide the handbook for their solution?

Congress writes back Reagan economies

Democrats prepare for battle

Washington

Members of the US Congress are digging in for a period of extended skirmishing over Mr Reagan's proposed cuts in federal support for research. Democrats are already fighting to restore many of the cuts, and in the House of Representatives at least — where they still have a majority — have already had some success. In contrast, Republicans are sticking as tightly as they can to the fiscal surgery recommended by the Office of Management and Budget (OMB).

Last week, the science and research subcommittee of the House Committee on Science and Technology voted to authorize a budget of \$1,161 million for the National Science Foundation (NSF) in the fiscal year 1982, 12 per cent more than Mr Reagan's proposed budget of \$1,033 million. As expected the committee voted to restore funding for science and engineering education. Mr Reagan's budget proposed that NSF involvement in these activities be terminated; but the subcommittee has restored \$75 million, close to the figure proposed by Mr Carter. It also agreed to restore \$26.1 million for research in the biological, behavioural and social sciences, another target of Mr Reagan's pruning.

Not all the subcommittee decisions went in the same direction. The members also approved a recommendation from Mr Allen Ertel (Democrat) to eliminate \$12 million which had been allocated for NSF's new Ocean Margin Drilling Program, initiated last year as a joint venture with the oil industry. A similar proposal to eliminate ocean margin drilling activities made by the House last year was overruled later by the Senate.

The subcommittee's recommendations on the NSF budget now go to the full science and technology committee, which will decide whether or not they should be passed intact to the floor of the House and where further opposition from Republican committee members is expected. The previous enthusiastic support for NSF by the full committee is now less certain — some Democrats may vote with the Republicans (as they have already done to push through proposed reductions in earthquake research and other projects sponsored by the Federal Emergency Management Agency).

In the Republican-dominated Senate it is generally expected that action on the NSF authorization will be held to Mr Reagan's level. If so, the major debate will come when the House and the Senate meet to

reconcile their approaches. These proceedings could be repeated when Congress decides how much money it wants to hand over to NSF rather than just to authorize.

A similar confrontation is brewing over the budget for energy research. President Reagan proposes major cuts in support for research into alternatives to oil and fossil fuels, such as solar and geothermal energy and alcohol fuels, arguing that these will be encouraged by proposed tax credits and that demonstration projects represent a hand-out to the private sector; in contrast he has recommended a substantial increase in funding for nuclear energy.

House Democrats are trying hard to reverse these proposals, sometimes with the assistance of Republicans who accept the economic rationality of federal investment in energy-saving technologies. The energy development subcommittee of the House Science and Technology Committee has already voted to increase the solar research budget from the \$193 million proposed by Mr Reagan to more than \$500 million, closer to Mr Carter's figures.

This proposal, if sustained by the full House of Representatives and if the Senate sticks with Mr Reagan's policy, is again likely to create substantial conflict. Aware of this possibility, the Department of Energy is said to be drawing up new legislation which would dramatically reduce its responsibility for developing new and renewable energy sources.

Despite Republican predictions of economic doom should the budget cuts not be approved by Congress — and threats of a presidential veto in this eventuality — many individual researchers remain confident that congressional allies will help save their projects. One example is research

in magnetohydrodynamics (MHD), aimed at improving the efficiency of coal-fired power stations, on which more than \$400 million of government funds has already been spent.

In its budget proposals the Reagan Administration suggests that research and development on MHD, previously scheduled to receive \$60 million in 1982, be terminated, arguing that the demonstration of engineering feasibility should become the responsibility of the private sector. A senior OMB official explained that the agency had "agonized greatly" over whether to make the cut; but that although MHD research met two of the government's present criteria for support — high-risk research with a long lead-time — it did not meet the third — a high probability that any successes would be taken up by private utilities.

MHD engineers, however, claim that many utilities remain committed to MHD, which is predicted to increase the efficiency of electricity production from coal by 50 per cent. They point out that cuts at this stage would not make sense, as most of the money is required to cover the operation of two plants whose construction is nearing completion, at the University of Tennessee's MHD Coal-Fired Flow Facility in Tullahoma, and the Component Development and Integration Facility for which superconducting magnets are being provided through the Massachusetts Institute of Technology.

Officials in Tennessee said last week that they expected local congressmen such as Representative Albert Gore or Senator Jim Sasser to argue successfully that the funds should be restored. In other years, their confidence might have been justified. This year, with OMB defining demonstration projects in energy research as "the new

Hidden cheer culled from statistics

Washington

In a neat display of the power of statistics, Reagan Administration officials have estimated that, even with the proposed cutbacks, spending on basic research will increase in real terms in the fiscal year 1982. Figures prepared by the Office of Management and Budget, assisted by staff from the Office of Science and Technology Policy, predict that funds for basic research will grow by 1.5 per cent from the 1981 level. At the National Science Foundation, where the cuts have focused on science and education and social and economic research, basic research in the natural sciences would show 3 per cent real growth.

The catch to these otherwise heartening figures is that they are based on an officially-predicted inflation rate of 7.5 per cent — an optimistic forecast that

assumes almost instantaneous success for Mr Reagan's economic policy.

Overall, the Office of Management and Budget expects the research and development budget for 1981 to be "a few hundred million dollars" higher than the previously expected figure of £35,600 million, largely because cancellations in civilian research will be more than offset by a \$500 million increase in military research. The net result of Mr Reagan's proposals would be to increase research spending by about 13 per cent between 1981 and 1982, to a total of \$40,600 million. Compared with Mr Carter's proposals, the research budget for the Department of Defense would be increased by an extra \$1,500 million, while the non-defence budget would be \$2,200 million less than first proposed.

David Dickson

pork barrel", there is more doubt.

Last week, for example, even the chairman of the House science and technology committee, Mr Donal Fuqua (Democrat) failed to persuade the members of the energy supply subcommittee to support his efforts to reinstate funds for a synthetic fuels demonstration project which faces shutdown without further federal support. Other areas in which efforts to restore funds are being made are for the Ocean Thermal Energy Conversion programmes — OTEC — described by OMB officials as being of very limited pay-off — and the decision to hold back on establishing a Center for Fusion Energy and its Fusion Energy Device, both centrepieces of Mr Carter's plans for the rapid development of magnetic confinement fusion.

David Dickson

UK aerospace

Eager for orders

British industry is beginning to wake up to space communications as a commercial proposition. British Aerospace, the largest British company with a space interest, is on tenterhooks, waiting to hear whether it has won two major contracts for communications satellites. The orders at stake are for a British defence communications satellite network and for the Arab regional satellite communications system, Arabsat.

The contracts would be a watershed for the company, for it has hitherto had only the European Space Agency as a customer for satellite systems. If its bids are now successful, it hopes to compete for similar orders in other parts of the world.

Arabsat, which is also being contested by Ford Aerospace and the Hughes Corporation, is the most crucial of the two. British Aerospace is cheered that it has appeared on the short list of three out of the five initial bidders. The Arab countries are to make a decision within four or five weeks.

Competition for the defence satellite system is not so fierce. The British government, which had been expected to award the contract at the end of March, is taking its time choosing between British Aerospace and Marconi. The plan is to order three satellites, worth a total of about £100 million, one of which will be operational at any time and which will replace the seven-year-old Skynet satellite. The company that wins will hope to gain a foothold in the NATO market.

British Aerospace's attempts to become a major supplier of telecommunications systems have also extended to setting up joint ventures with other companies. Together with Plessey, which has built satellite earth stations, it says it can offer a team capable of building complete defence communications satellite systems. It is also planning a joint venture with the US Comsat General corporation to lease military satellite communications services. And its ambition to compete for more

contracts like Arabsat led it last month to set up a group called Satcom International with the French company Matra. Satcom plans to put in a bid for a regional satellite for Australia in May.

British Aerospace is selling itself on its experience. It has played a leading part in the telecommunications activities of the European Space Agency, having been the prime contractor for the series of Orbital Test Satellites and the more sophisticated European Communications Satellites. It also prides itself on being the only non-American company to contribute to Comsat General's Comstar programme for long distance telephony within the United States.

Attempts to exploit the commercial potential of space may have come rather late in Britain. But the present flurry of activity is not confined to aerospace and electronics companies. The Department of Industry says that an increasing variety of organizations, from banks to pension funds, now think of telecommunications when asking where to put their money.

How matters will develop depends not only on the industry's ability to compete and fulfil its promises but also on what the British government decides. The industry is now in limbo waiting for a government response to a confidential report by the Central Policy Review Staff on the commercial exploitation of space and for the publication of a Home Office report on direct broadcast television by satellite. The reports, if nothing else, may recommend making it easier for venture capital and private enterprise to take a share of the action.

Judy Redfearn

UK nuclear power

Battle begun

The first shots have been fired in the impending battle about the building of a pressurized water reactor (PWR) in Britain, even though the Department of Energy has not yet decided when the proposed public inquiry will be opened. Last week, the Friends of the Earth (UK branch) issued a document subtitled "A critique of the government's nuclear power programme" (Friends of the Earth, London; £1.50) which concludes that its

own case against the proposed reactor, to be built at Sizewell in Suffolk, on the North Sea coast, is so strong that a public inquiry is not necessary.

The argument is straightforward. First, forecasts of electricity demand in the United Kingdom, revised downwards for the past several years, are "wholly unrealistic" and are likely to be further undermined by changing patterns in the demand for electricity. Second, the cost of building the first British pressurized water reactor is likely to be twice the Electricity Generating Board's estimate (of £850 per kilowatt) while the cost of generating electricity will be magnified by the alleged low availability of reactors based on the Westinghouse design.

Third, the report (which includes a helpful account of the design changes required by the British nuclear inspectors) says that "the unique hazards of the American PWR must not be introduced into Britain". It is especially critical of a study by the nuclear inspectors of the generic safety of this type of reactor, which has not been published in full. The report also argues that "unresolved" problems concerning the safety of these reactors have not yet been resolved.

The conclusion is that Britain (and British taxpayers) would be better advised to put their money into renewable resources. Insulating lofts to a high standard would be six times as productive as building a pressurized water reactor at Sizewell, according to the Friends of the Earth.

Earlier this week, the Central Electricity Generating Board — the chief but not the only target of the Friends of the Earth — had not decided whether, and if so how, to reply to the criticisms levelled at it. To the extent that some of the complaints coincide with the criticisms of the House of Commons Select Committee on Energy some weeks ago (see *Nature* 19 February p.621), the authors of the new report may have to wait until the formal response from the Department of Energy.

The intended public inquiry about the proposal to build a single pressurized water reactor at Sizewell is now a formal commitment by the British government. Although the form of the inquiry has not yet been settled it is likely to be generic, dealing not with the characteristics of the Sizewell site

Magnox at Dungeness A — open soon?



but with the virtues (or lack thereof) of pressurized water reactors in general. Most probably, the inquiry will not now begin until the second half of 1982, implying that construction of the reactor system cannot begin until late in 1983.

For the electricity undertakings, the chief objective of the proposal to adapt an American design to British circumstances is the hope that pressurized water reactors could be built more quickly than indigenous gas-cooled designs. The Friends of the Earth document, however, is sceptical of the six-year construction time on which the generating board is counting.

Meanwhile, the Central Electricity Generating Board is somewhat encouraged that four of its oldest gas-cooled reactors, shut down some months ago after the discovery of cracks in the expansion joints of the gas-cooling circuits, can soon be started up again. The first of these reactors, at Berkeley on the Severn, is to be restarted in a few months following a demonstration that the cracks have not grown since the reactors were first built.

Czech nuclear power

Action on safety

Czechoslovakia's nuclear energy authorities have recently made overtures to their British counterparts, suggesting that the agreement on cooperation in nuclear safety, signed almost 15 years ago, should be revitalized. In particular, the Czechs are seeking an exchange of information and expertise on the disposal of nuclear waste and on the training of power station personnel.

The human factor in nuclear accidents has lately become a major concern in Czechoslovakia. Last August, after repeated denials, and without any attempts to prepare public opinion, the authorities suddenly admitted that in the past few years there had been two major accidents at the Jaslovské Bohunice nuclear power station.

During a recent phone-in programme on Prague radio, Dr E. Medke, director of the Nuclear Power Station Research Institute at Jaslovské Bohunice, described the institute's "high quality" research programme, aimed at the "long-term, safe, reliable and in particular, economically effective" operation of Czechoslovakia's nuclear power stations. This, he said, is to include work on the human element in safety. Analysis of nuclear accidents throughout the world, he said, shows that only 20 per cent were caused by equipment failure and the rest by operator error.

A training programme is under way to provide the 1,300 graduate engineers whom the country's nuclear programme will need by 1990. After taking a first degree in engineering, the candidate will go on to a two-year theoretical course, and then to a power station for practical

training. This will include work on a simulator, now being built at Trnava. After careful screening, the student will then become an assistant at a nuclear power station, and will finally take examinations set by the State Commission to qualify for a nuclear engineer's licence.

Czechoslovakia, which is more than self-sufficient in uranium, has a major commitment to nuclear energy, and accordingly tried to hush up the accidents at Jaslovské Bohunice, one of which involved two deaths. But in 1978 the Charter-77 movement published a detailed report on the incidents — and although then still unprepared to acknowledge that anything untoward had occurred, the authorities immediately inaugurated a policy of recruiting nuclear power station workers of a better psychological calibre — a major recommendation in the Chartists' report.

In the radio programme, Dr Medke made no mention of the accidents, and left his deputy, Dr S. Novak, to deal with public apprehensions about nuclear power. Regarding the station now under construction at Mochovce near the Hungarian frontier, Dr Novak said that there has never been a case in which one country has contaminated another's territory by nuclear affluent, and that all Czechoslovak power stations have adequate safety measures to prevent environmental damage in the case of breakdown or failure. This statement contradicts the Chartists' report which alleges that in one of the incidents at Jaslovské Bohunice, effluent reached a watercourse. It also contradicts claims from Bulgarian environmentalists that radioactive contaminants can be found in the Danube — although these are conventionally assumed to be of "capitalist" origin.

Vera Rich

Soviet space programme

Mongolia's turn

Last week's spaceflight of the Mongolian cosmonaut, Jugderdemidyn Gurragchaa, coincided with the start of the diamond jubilee celebrations of the Mongolian People's Republic. A few days before the launch, Mongolia's new Five Year Plan was published, extolling the value of science and technology to the development of the republic.

Not surprisingly, most of the experiments performed by Gurragchaa and his Soviet colleagues aboard the Salyut-6 were aimed at practical applications rather than basic research. One of the major tasks was the mapping of Mongolian natural resources, using conventional cameras and a Bulgarian "Spektr-15" multichannel spectrophotometer. This survey will provide data for existing joint Comecon research projects on Mongolian resources including the irrigation of the Gobi Desert. A programme of simultaneous photo-



Gurragchaa — applying science?

graphy of the Salyut spacecraft from Mongolia's two satellite tracking stations (in Ulan Bator and the Gobi) was intended to provide data on factors affecting space geodesy.

The biological programme involved the monitoring of the cosmonauts' biorhythms, work capacity, cardiovascular function and psychological response to the conditions of space travel. The biorhythm observations were started in 1977 during the selection of candidates for cosmonaut training — apparently the first time that a psychological and biochemical study of Mongolian citizens had ever been undertaken on a daily basis. Technological experiments included the study of natural quartz and of various alloys of copper under zero gravity.

Fundamental research was not completely ignored in Gurragchaa's mission. An experiment codenamed in Russian "Izlučenje" (radiation) continued the Mongolian work on primary cosmic rays with energies in the TeV range which was begun in 1972 with an experiment designed by S. Sugar and B. Chadraa of the Laboratory of High Energy Physics of the Academy's Institute of Physics for the joint Comecon Interkosmos-6 satellite in 1972.

Vera Rich

UK computer industry

ICL falters

International Computers Limited (ICL), Britain's largest computer manufacturer, announced last month that it has been trading at a substantial loss since the middle of 1980. The news of ICL's problems came only a few weeks after similar news from Imperial Chemical Industries (ICI). Until the beginning of this year, both companies had been regarded as profitable examples of Britain's otherwise ailing industry.

Unlike ICI, however, ICL is to be bailed out by government money. Sir Keith Joseph, Secretary of State for Industry, has said that the government will act as guarantor for up to £200 million worth of

loans which the company is expected to secure from the banks. Help is also to be provided for ICL's research and development programme which last year consumed about 8 per cent of the company's £715 million turnover.

Sir Keith's intervention is seen as a major blow to the Conservative government's policy of not interfering in private industry. But the decision was taken only after several months of searching for private investors had found no takers. The chief justification for supporting ICL when other private industries are allowed to go to the wall seems to be that the smooth operation of many government offices depends heavily on its computers.

ICL is something of a favourite of British governments. When it was created in 1968 from the amalgamation of two computer companies, the government retained a substantial shareholding and supported it heavily during its early years both with cash and with awards of almost all contracts for large computers for government departments. In the 1970s, the company became profitable and in late 1979, Mrs Thatcher's government sold off its shares. But only 18 months later, ICL's profits took a lurch.

Precisely why the company's fortunes changed so suddenly seems to be a mystery. The company itself says that the world recession combined with the strong value of the pound and high interest rates is responsible. There were hints of trouble, it says, in the first half of 1980 but during the second half both home and overseas orders fell off markedly. Others have said that poor management, a concentration on large computers when the market is moving towards desk-top machines, spending too much on parts of the business not connected with manufacturing and even underpricing goods, are also responsible.

Neither the government nor the banks, however, seem to have stipulated conditions for the £200 million loan. The company hopes to ride out the recession by continuing business much as before and economizing wherever possible. It says, for example, that its research and development programme will continue at much the same level as last year, with the same mix of research on large and small machines. One of the projects on which it pins hopes for the future is the large Distributed Array Processor which has so far been sold to the British National Oil Corporation, for seabed exploration, and Queen Mary College University of London. The Science Research Council is considering buying one for its Rutherford or Daresbury laboratories.

Sir Keith Joseph hopes that the government will not have to make good its guarantee, believing that the loan will see ICL through the recession until business picks up again. Meanwhile, the net is still being cast far and wide in the hope of finding private investment on suitable terms.

Judy Redfearn

London medical schools

Mergers galore

The University of London has made a modest beginning on the reorganization of its twelve medical schools. Last week, the senate of the university approved a plan to save something like £5 million a year, or ten per cent of the present cost to the university of medical education. Some two-thirds of the saving will come from a general reduction of staff-student ratios throughout the medical schools.

The accompanying reorganization of the London medical schools is only a pale shadow of the grand design put forward more than a year ago by a committee under Lord Flowers, rector of Imperial College. At the senate meeting last week, it was formally agreed that the Westminster Hospital Medical School (a clinical school now recruiting about 100 students a year) should be merged with the medical school at Charing Cross Hospital (now located in West London). Other proposed groupings of medical schools have been given until later in the year to work out how they will coordinate their teaching.

The Westminster school considers itself unfairly singled out by these developments. Ever since the publication of the Flowers Report, Westminster has been urging the benefits of small medical schools, complaining that its students or their successors will be lost in a medical school with an annual intake of 180, the planning figure for clinical students at Charing Cross. It also says that the efforts that have been made in the past year or so to coordinate teaching between the two hospitals may be jeopardized by the shotgun marriage now decreed. Given forward commitments to students, it may be as much as five years before teaching comes to an end on the Westminster site.

In retrospect, the schools least affected by the changes now proposed are those at Guy's Hospital and St Thomas's Hospital, which announced last year that they would form a joint school, with a single governing body, but would continue to operate on both the present sites. This arrangement has now been sanctioned by the university senate. The other consortia from which the university is seeking financial savings are University College, Middlesex and St Mary's, which are required between them to save the equivalent of one clinical and one preclinical school; St Bartholomew's and the London Hospital, which have been given until June to suggest how they will merge; and King's College (which has a preclinical school) and King's College Hospital, which seem to be willing to fall in with the recommendation that they should form an integrated unit.

The proposals accepted last week by the senate were recommended by the Joint Planning Committee on Medical Education, set up to pilot the Flowers Report through the university. In the past

year, and especially as a result of a detailed analysis of the costs of the London medical schools carried out by a firm of accountants, it became clear that changes of staff-student ratios offered the most rapid route to cost reductions. The committee accepts as a standard for future staffing a ratio of 1:10 for preclinical schools and 1:7 for clinical schools. This would mean the loss of 180 academic posts, which it is hoped can be achieved by natural retirement and wastage in the next five or six years.

Running through the report of the joint committee is the view that medical education should be concentrated in larger units than at present, especially where modern buildings have been provided. In this spirit, the committee recommends that the development of the St George's Hospital site in South London should be continued. The future of the postgraduate medical institutes remains entirely open, however. With the exception of the Institute of Dermatology (which is to merge with St Thomas's Hospital Medical School), the institutes have been seriously affected by the reductions in the past year in the numbers of overseas students. The university is hoping that central government will help bail them out, but no decision has yet been reached.

Last week's decision by the university senate is likely to be regarded, within the University of London, as some kind of precedent for the reorganization of the non-medical parts of the university likely to be recommended by the Swinnerton-Dyer committee, due to report in September, and elsewhere as a sign that, if pressed hard, even British universities can change their ways.

Law of the Sea

US backtracks

Washington

The Reagan Administration's decision to carry out a complete review of the US position on the draft Law of the Sea Treaty, now under discussion at the United Nations in New York, indicates that significant new attitudes may emerge towards international resource disputes ranging from the surface of the Moon to the minerals of the Antarctic continent.

The United States is considering whether to adhere to the concept that natural resources which lie outside the immediate political domain of individual nations should be treated as "the common heritage of mankind" — a term coined by the United Nations — or whether, in the name of free enterprise and the national security implications of a shortage of strategic minerals such as cobalt and manganese, this concept should be abandoned, together with attempts to set up international regulation to control exploitation.

A new twist to the Law of the Sea negotiations comes from the likely take-over of the Kennecott Corporation, a leading critic

of the United States' negotiating position, by Standard Oil of Ohio (Sohio). Sohio is 53 per cent owned by British Petroleum, which in turn is 45 per cent owned by the British government. In the past, BP Minerals, a member with Kennecott of one of the five consortia interested in deep-sea mining, has been keen to finalize the treaty to protect its existing exploration and mining operations.

In theory, a Sohio/Kennecott deal could favour a successful outcome to the Law of the Sea negotiations. In practice, the outcome could depend more on whether the Reagan Administration listens to arguments from international companies that a treaty is essential to the security of their future trading operations, or whether it gives greater weight to conservative pressure groups arguing that the whole "common heritage" concept, along with Third World demands for a "new international economic order", should be dismissed as socialist subversion.

In dismissing several key members of the US negotiating team, and initiating a complete review of the treaty in its draft form, the Reagan Administration is responding to complaints from groups such as the Council on Economics and National Security. Council officials were among the principal authors of a warning by the American Geological Institute to the two presidential candidates last September of a Soviet "resource war" aimed at depriving the United States of materials indispensable to its national defence.

Immediately after the election, the Reagan Administration set up a task force on national strategic minerals policy, which recommended a complete review of the draft Law of the Sea Treaty. One of the principal opponents of the treaty in its present form is a Washington attorney, Leigh Ratiner, acting head of the Department of the Interior's office of oceans and minerals under President Ford, and later the chief lobbyist on the treaty for the Kennecott Corporation.

Mr Ratiner is acting as counsel to the L-5 society, a 4,000-strong group of scientists and "futurists". The society is largely credited with stalling any effort by the Senate to ratify a draft treaty covering mining operations on the Moon, agreed in principle by the United Nations last year.

Particularly strong opposition to the treaty has been voiced by the US Steel Corporation of Pittsburgh, a leading partner in the Ocean Mining Associates consortium. According to the chairman of US Steel, William R. Roche, policies pursued by the Carter Administration tended to increase the country's strategic mineral vulnerability.

Attempts to scuttle the Law of the Sea negotiations have been fiercely attacked by Mr Carter's chief negotiator Mr Eliot Richardson, ex-ambassador to the United Kingdom, who warned that many multinational companies would face severe consequences if they tried to disrupt

agreements that had already been reached by US negotiators in good faith.

Many conservatives, however, see Mr Richardson's position as symptomatic of a "soft" approach to the Third World that does little to protect US economic and strategic interests. A change in strategy on the Law of the Sea Conference could also have implications for other negotiations.

Following the Antarctic Treaty of 1959, the 12 (now 14) nations laying claim to Antarctic territory have agreed voluntarily to restrain mining activities until an appropriate international treaty can be worked out. Yet this too is now being challenged by conservatives within the Reagan Administration, largely on the grounds that negotiations could result in the United States gaining less than it might expect from an international free-for-all, given its superiority in mining technology.

Delegates attending the Law of the Sea negotiations session in New York, although dismayed with the Reagan Administration's position, have agreed to

discuss the details of the draft treaty. Some had hoped that the Kennecott/Sohio deal, together with the influence of other international groups such as Royal Dutch Shell, might help to dilute the pressure of more nationally-oriented companies on the Reagan Administration.

Meanwhile, a group of lawyers, scientists, environmentalists and economists has issued a report calling for research into the character of the marine environment of the deep ocean, and its susceptibility to mining operations. The group points out that no explicit mention is made of the need for prior environmental assessment either in a bill on deep-sea mining rights passed recently in West Germany, or in draft bills now before French and British legislatures. It was also concerned that such nations might get together with the United States and other countries to set up their own institution for regulating mining activities before the Law of the Sea negotiations finally break down.

David Dickson

India mines the sea

New Delhi

India has become the first developing country to collect polymetallic nodules from the ocean floor, a distinction achieved by India's first research ship, the *Gaveshani*. The 1,900-tonne vessel, once a hopper barge and converted into a research ship in 1975, is operated by the National Institute of Oceanography (NIO) in Goa, one of the 40 institutes under the aegis of the Council of Scientific and Industrial Research.

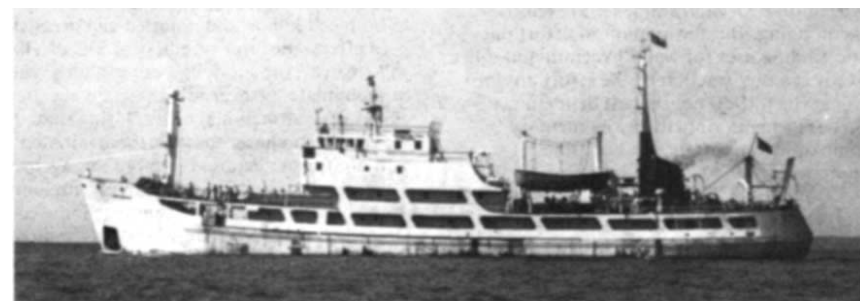
NIO has been preparing the *Gaveshani* for nodule hunting since 1978 and the catch was made in January during her cruise from Goa to Mauritius, which was led by Dr S.Z. Qasim, director of NIO.

According to Dr Qasim, the *Gaveshani* located a nodule-rich region in the Indian Ocean about 700 miles from the mainland and reaching almost as far as Mauritius. During the return journey the ship collected more nodules from several locations at depths of 3,600–5,400 metres.

Besides nickel, cobalt, copper, manganese and iron, the samples are believed to contain traces of gold, but their precise metallic contents are now being analysed by eight Indian laboratories.

The *Gaveshani*'s nodule collecting gear

Gaveshani — hunting for nodules



consisted of a 10,000-metre winch, underwater camera, dredges and some twenty "boomerang grabs", which go down to pick up the nodules and come up automatically. According to Dr Qasim, the surveyed seabed area had a nodule concentration of 1 to 5 kilogrammes per square metre, which is potentially economic for mining.

India is not ready to start mining proper, but NIO plans to map the nodule-rich region to assess its total mineral wealth. The operation should be speeded up next year when the *Gaveshani* is joined by a second research ship now under construction in West Germany. Since India imports all its nickel and cobalt and more than 60 per cent of its copper requirements, the discovery of a nodule basin in its close vicinity should be welcome news.

There is a long haul between picking up a few seabed nodules and being able to mine them commercially. Yet the *Gaveshani*'s demonstration has raised some eyebrows among nations that at present have a monopoly on ocean mining. It is likely that United Nations talks on the Law of the Sea will acquire a new dimension as a result of India's unexpected entry into the ocean mining club.

K.S. Jayaraman

CORRESPONDENCE

Licking Darwin

SIR — How absurd to hold that Darwinism cannot be falsified. All it would take is a single clear demonstration of the *de novo* creation of a complex organism, well adapted to its environment, and the Darwinists would be legless.

G.P. VINSON

Department of Biochemistry and Chemistry,
The Medical College of St Bartholomew's
Hospital, London EC1, UK

Weed control

SIR — As May emphasizes¹, all possible steps should be taken to prevent the intrusion of noxious weeds into new regions. That article paid rather less adequate attention to weeds that are already established. Consider two hypothetical situations. Suppose it was discovered that one of these weeds contained a valuable fibre. Within a few years we would have a conservation programme to save the weed from extinction. Suppose enterprising adolescents found that one of the weeds contained a hallucinogen. The \$15 million quoted as the annual cost of water hyacinth control in Florida and Louisiana would be "chicken feed" compared with what drug control authorities would spend on elimination.

Water hyacinth has been studied for so long that components as dramatic as these are unlikely to be found in it. Nevertheless, it is a prodigious producer of biomass and is often conveniently located on broad areas of water from which mechanical collection is easy. Twenty one years ago² I made "... a plea for some reorientation of thinking in areas suffering from, or threatened with, water hyacinth. There is a case for directing some research towards finding uses for the pest instead of concentrating all efforts on its eradication."

Since then, the pest has made great headway. Less headway has been made on finding uses for it. Interest is, however, increasing. Chibbar and Singh³ fed cows on silage made from water hyacinth and paddy straw, and there have been several recent papers in the *Indian Journal of Animal Science* extending that work. Hyacinth is deliberately cultivated in China for pig feed. It is a component of the mixtures used for biogas fermentation in many countries. These and other uses are surveyed in *Making Aquatic Weeds Useful* (US National Academy of Sciences, 1976) and in *Handbook of Utilization of Aquatic Plants* (Food and Agricultural Organization, 1979). This is encouraging. But the amount of effort put into finding uses for water hyacinth and other weeds is a tiny fraction of the effort put into killing them. I do not suggest deliberate encouragement, only the assignment of comparable importance to both lines of approach.

N.W. PIRIE

Harpenden, Herts, UK

1. May, R.M. *Nature* **290**, 85-86 (1981).

2. Pirie, N.W. *Nature* **185**, 116 (1960).

3. Chibbar, S.S. & Singh, G.D. *Indian Farming* **20**, 24 (1971).

Cafeteria feeding

SIR — I wonder if I am alone in my dislike of the term "cafeteria feeding" used first, I think, by Rothwell and Stock in 1979 (Rothwell, H.J. & Stock, M.J. *Nature* **281**, 31; 1979). According to the *Concise Oxford Dictionary* a cafeteria is a "restaurant in which customers fetch food and drink from a counter" which is hardly an accurate description of the behaviour of rats allowed to make their own choice of palatable items of food. I would like to suggest that, before it is too late, editors of scientific journals ask authors to use a more accurate description, for example to replace "cafeteria diet" by "a self-selected palatable diet" (abbreviated SSPD) or other equivalent term.

A.S. COLE

University of Bristol, UK

Cancer causation

SIR — The recent series of articles¹⁻³ on cancer causation prompts me to respond. Without defending Epstein and Swartz, I would like to offer several alternative explanations to some of the examples which Cairns used to argue his case that somatic mutations play little or no part in human carcinogenesis.

Although Cairns' novel hypothesis is to be welcomed as a stimulus to new experimental approaches, I feel that he has summarily dismissed a significant amount of evidence consistent with the somatic mutations theory of cancer. Furthermore, carcinogenesis is a multi-step process and there are probably several distinct molecular mechanisms that can "cause" and influence carcinogenesis, not the least of which are somatic mutations⁴.

Specifically, Cairns argues that the apparent lack of common fatal cancers in xeroderma pigmentosum (XP) individuals (the cells from whom lack proper DNA repair causing them to be hypermutable to many mutagens, including the ultraviolet component of sunlight) is evidence that the production of mutations in XP cells is probably irrelevant to carcinogenesis. There are at least two reasonable alternative explanations why XP patients rarely get cancers other than skin tumours⁵. The XP patients normally do not live long enough to manifest these internal tumours (an explanation cited but dismissed by Cairns) and most cancers are the result of initiation (mutagenesis) and promotion (amplification of an initiated cell into a clone where additional genetic or epigenetic changes can occur). One can explain the production of skin tumours in XP on this hypothesis since, using this concept, repaired sunlight-induced lesions in the skin cells of XP patients will lead to both cell killing and mutated viable cells.

In effect, the large numbers of killed cells will "force" the surviving, but initiated, cells to repopulate the necrotic tissue, giving rise to clones of dysfunctional cells. During that repopulation phase, some of these initiated cells are further exposed to more mutagens. Thus, the cytotoxicity caused by the inability to repair large amounts of sunlight-induced DNA damage can be considered an "indirect promoter", much in the manner of a partial hepatectomy⁶. On the other hand, internal cells of XP patients, which are exposed to

chemical mutagens that mimic sunlight-induced DNA damage, probably are not exposed to high enough levels constantly to cause significant cell killing, as occurs in the skin. In effect, these initiated cells will not be promoted! Only after long periods of time or exposure to chronic effective levels of exogenous or endogenous tumour promoters will internal tumours appear.

A second argument used by Cairns to support this thesis of genetic transposition was the discussion on Bloom's syndrome. I agree with Cairns that Bloom's syndrome might be a "better" model for studying carcinogenesis, but for a different reason. At the time of writing his paper, Cairns was not aware of the observation that Bloom's syndrome seems to be a "mutator" syndrome⁶, in that it spontaneously generates mutations in their somatic tissue. In effect, Bloom's cells can generate mutations in any proliferation cell of the body, with or without exposure to "environmental" mutagens. This would explain the wide range of tumours (and other clinical features) associated with the syndrome, and would also fit the initiation/promotion model, since each time a spontaneous mutant cell arises and then divides, it can enhance the probability of additional genetic imbalances.

JAMES E. TROSKO

Michigan State University,
East Lansing, Michigan, USA

1. Epstein, S.S. & Swartz, J.B. *Nature* **289**, 127-130 (1981).

2. Cairns, J. *Nature* **289**, 353-357 (1981).

3. Editorial *Nature* **289**, 431-432 (1981).

4. Trosko, J.E. & Chang, C.C. in *Chemical Carcinogens and DNA Vol. 2* (ed. Grover, P.L.) 181-200 (CRC Press, Boca Raton, Florida, 1979).

5. Trosko, J.E. & Chang, C.C. *Med. Hypothesis* **6**, 455-468 (1980).

6. Warren, S.T. et al. *Proc. natn. Acad. Sci. USA*. (in the press).

Genes and free will

SIR — Asked to dissociate themselves from the neo-Nazi use of their names and theories, John Maynard Smith¹ and Richard Dawkins² have both done so. Professor Maynard Smith's unequivocal couple of sentences will be no surprise to scientists who know him as a man of principle and integrity; his letter will give no comfort to the National Front or any similar grouping.

It is a pity, though, that Dr Dawkins appears to think that the issue I sought to raise in my letter is of a parallel with the relatively harmless debate on cladistics and dialectics which has kept us entertained over recent months. If it were not for that dreary little edge of "reds under the bed" of which Stephen Jay Gould reminds us³, I suspect we could look forward to *Nature* with the same enthusiasm as once we had for *Beano*. However, the issues are not the same. Racism directly haunts the streets of our cities. In our laboratories neither Dawkins nor I live in the daily hazard that the obscene rubbish of the *New Nation* and its followers fosters in the East End or New Cross. When Dawkins disdains the ephemera of mere "human politics", or accuses his critics of being as guilty as the Nazis of dragging the elegant purity of his neo-Darwinism into the litter of the city streets, he does so at the peril of a repetition of the tragedies of the 1930s.

Of course we must all beware of the fallacy

Continued on page 432

NEWS AND VIEWS

Nerves, myelin and multiple sclerosis

from Miranda Robertson

At the beginning of December, a select group of clinicians came together with a similar group of research biologists in Berlin to plumb the depths of their respective ignorance about multiple sclerosis and the glial cells of the central nervous system. Multiple sclerosis is baffling in almost every respect — its cause is unknown, its pattern erratic and its course extraordinarily variable. The target of the disease is the myelin that insulates the nerve fibres of the central nervous system, or the oligodendrocytes that produce it. The symptoms are thus due to blocked or unreliable conduction along nerves, but it is not known what induces the attack on the myelin; nor is it understood why, rather than steadily worsening, the symptoms should alternately flare up and partly subside over the progressive course of the disease; nor why some patients may remain free of symptoms for years after a first attack, whereas others will suffer repeated relapses and die within a decade.

Two of the principal aims of the meeting*, whose main topic was the relationship between neurones and glia but which also included much discussion on viral and immunological mechanisms, were to try to establish, first, what is going on in multiple sclerosis patients, and second, what might be done about it.

Probably the most cogent view of the characteristics of the disease is due to the immunologists, whose contribution has recently been discussed in *Nature* by Bloom¹. The pattern of exacerbation and remission of symptoms, for example, may represent defective (and thus irregularly oscillating) regulation of an immune response, manifesting itself as recurring episodes of inflammation — an explanation consistent with the efficacy of steroids during acute attacks (B. Arnason, Chicago University).

The cause of the inflammation and tissue

destruction could be a latent virus infection and the immune response to it, or an autoimmune attack on the central myelin that might or might not have been stimulated in the first instance by viral infection. In animals, the pathological changes can be mimicked by some viral diseases, or by immunization with central myelin — though in the latter case, the disease follows a chronic relapsing course only in animals with genetic immune deficiencies. (This may be significant in the light of evidence in man for linkage with specific HLA antigens: see below.)

However, multiple sclerosis itself is not associated with a specific immune response either to a particular virus or to a component of the central nervous system; nor is there any evidence that immunosuppressive treatment can alleviate acute attacks of the disease. This raises two questions. The first is that of the relevance of any of the existing animal models to the human disease. The second is whether there is any point in looking for immune correlates of multiple sclerosis without having identified a specific antigenic target: immune activity in itself is a very poor guide to aetiology, since tissue degeneration from any cause (including, for example, stroke) will attract the attention of the immune system.

Nonetheless, many laboratories have looked for changes in the immune response that may be associated with multiple sclerosis, and one of the most recent fruits of such research has been the discovery that exacerbations of the disease coincide with a reduction in the level of circulating immunoregulatory T cells². This may mean that the failure of immune regulation unleashes the attacks (in which case, however, why specific to central myelin?), or alternatively that suppressor T cells and oligodendrocytes share an antigen and the disappearance of the immunoregulatory lymphocytes is a concomitant and not a cause of the attack. Many participants at the meeting took the view that the latter was the more interesting possibility

because it could lead to the identification of the antigenic target — an advance whose importance it would be hard to exaggerate.

There is, of course, a strong possibility that there is no single cause of multiple sclerosis. It was generally agreed, for example, that if the disease is a response to viral infection it is very likely that more than one virus can elicit it. But some very preliminary genetic evidence suggests that the heterogeneity of multiple sclerosis may run deeper than that. Nowadays, if nothing else is known about a disease — in fact, particularly if nothing else is known — its HLA association is. Multiple sclerosis in Europe is in linkage disequilibrium with HLA antigens D2 and D4 (ref. 3), but more recent investigations have revealed a further link, with DR3. The significance of these associations is unknown, but their weakness, together with genetic evidence from family studies, and in particular from twins, strongly suggest that HLA D2 and D4 are in linkage disequilibrium with some other, unknown genetic factor predisposing to the disease (W.I. McDonald, London University). The case of DW3, on the other hand, may be different: DW3 is associated with two known autoimmune diseases, myasthenia gravis and autoimmune thyroiditis; there is thus a possibility that its prevalence reflects a defect in immune regulation predisposing to autoimmune attack — the exact target presumably depending on other factors.

If the ultimate cause of multiple sclerosis remains unknown, can anything be done in the meantime about its proximal cause? Can conduction be restored in demyelinated nerves, or dead oligodendrocytes be replaced? Drugs that enhance conduction have been tried, but without much success. One of the most recent efforts in that direction has been with 4-aminopyridine, which blocks potassium channels and prolongs the action potential⁴. This facilitates

Miranda Robertson is an associate editor of *Nature*.

*The Dahlem Conference on 'Neuronal-Glia Cell Interrelationships' was held in Berlin on November 30–December 5 1980. A Dahlem Workshop Report on the conference will be published.

conduction but unfortunately at the same time promotes seizures. It is, in any case, doubtful how far simply overcoming conduction block could restore useful function, since both the pattern and the spread of impulses would inevitably be grossly abnormal.

Some natural short-term recovery of function, on the other hand, is to be expected on the basis of recent investigations on the distribution of sodium channels in myelinated nerves. In the peripheral nerve, sodium channels are clustered at the nodes of Ranvier, being scarce or absent in the membrane under the myelin sheath⁵. This means that immediately after the demyelination of a peripheral nerve, conduction is completely blocked and it is not until some days afterwards that sodium channels begin to appear at the internode region and action potentials can be propagated across it (T.A. Sears, London University). If the same differential distribution of channels holds in the central nervous system (and there is some evidence that it does⁶), the exacerbating and remitting pattern of multiple sclerosis may partly reflect the

length of time it takes for sodium channels to redistribute themselves after demyelination.

Real recovery could be brought about only by the proliferation of oligodendrocytes and remyelination of axons. Although Schwann cells are known to proliferate and remyelinate peripheral fibres, and central remyelination commonly occurs in other demyelinating diseases, it does not seem to occur in multiple sclerosis. This suggests that multiple sclerosis may destroy the oligodendrocytes which, unlike Schwann cells, cannot be replaced. This reflects the more advanced state of differentiation of the oligodendrocyte. Not only can Schwann cells in culture be induced to proliferate, both by contact with axonal membrane (R.P. Bunge, Washington University) and by a soluble pituitary protein⁷, but they can also be induced by co-culture with nerves to increase 1,000-fold the synthesis of the principal myelin protein; or, by co-culture with muscle cells, to synthesize acetylcholine. Cultured oligodendrocytes, on the other hand, will synthesize the principal central myelin proteins in the absence of

neurons, although they do seem to depend on astrocytes to support proliferation (R. Mirsky, London University). What signals either Schwann cells or oligodendrocytes *in vivo* is quite unknown, although with the recent development of specific markers for central and peripheral neurones and glia, this question is now open to investigation *in vitro*.

Whether such investigations will make any direct contribution to the treatment of multiple sclerosis, or whether the capacity for recovery in the central nervous system lies in its inherent plasticity rather than in regeneration, only more research can show. In the meantime, everyone agreed that prevention is always preferable to cure; in the case of multiple sclerosis, it may be the easier to achieve. □

1. Bloom, B. *Nature* **287**, 275 (1980).
2. Antel, J.P., Richman, D.P., Medof, W.E. & Arnason, B.G.W. *Neurology* **28**, 106 (1978).
3. Batchelor, J.R., Compston, A. & McDonald, W.I. *Br. med. Bull.* **34**, 279 (1978).
4. Sherratt, R.M., Bostock, H. & Sears, T.A. *Nature* **283**, 570 (1980).
5. Chiu, S.Y. & Ritchie, J.M. *Nature* **284**, 170 (1980).
6. Kocsis, J.D. & Waxman, S.G. *Nature* **287**, 349 (1980).
7. Brockes, J.P., Lemke, G.E. & Balzer, D.R. Jr *J. biol. Chem.* **255**, 8374 (1980).

A new model for phototransduction

from Peter G. Lillywhite

IN this issue of *Nature*, R. Payne and J. Howard (see page 415) put forward a new model for phototransduction which provides an elegant alternative to the cascade model of Hodgkin¹. A successful model for phototransduction in vertebrates and invertebrates needs to explain (1) the latency of the response, (2) the increase in sensitivity to light steps and flashes with dark adaptation, and (3) that large increases in sensitivity can occur with only small changes of time scale. In the model invented by Hodgkin and fitted to *Limulus* receptors in collaboration with Fuortes¹, a cascade of stages, each occurring with an exponential delay, was assumed. Several stages are required because a single step would not, for example, exhibit latency. It turned out that about ten exponential stages were required to mimic the responses of *Limulus* receptors. Strong illumination was assumed to decrease the time constants of most, but not all of the exponential stages of transduction, causing the decrease in sensitivity observed with light adaptation.

It was immediately attractive to give physical reality to the hypothetical chain of exponential stages or 'events' as chemical reactions. Borsellino *et al.* showed in 1965 that, if it is assumed that the time constants of the exponential delays represent the rates of chemical reactions, the effect of temperature upon voltage response fits the classical Arrhenius relationship². Although they took the necessary precaution of pointing out that consistency

with the Arrhenius relation was not proof that the hypothetical chain of chemical reactions actually existed, chemical reaction chain models are nevertheless highly popular. For example, Baylor *et al.* recently (1974) modelled the responses of cones in the turtle retina using a chemical reaction cascade model³.

Payne and Howard now propose an alternative mathematical and conceptual model to that of Fuortes and Hodgkin and other similar cascade models. The model is attractive, because it describes locus photoreceptor responses (which are very similar to those of *Limulus* modelled by Fuortes and Hodgkin) almost perfectly, accounting for the three characteristics listed above, but does not require a chain or cascade of many hypothetical stages. The cascade models of Fuortes and Hodgkin and Baylor *et al.* have an unattractive side to them, which is the large number of implied transduction steps or reactions. To fit the photoreceptor responses they studied, Fuortes and Hodgkin required about ten steps and Baylor *et al.* required nine (six for creation and three for destruction of ion channel blocking particles). When Payne and Howard tried to use these models to fit locust photoreceptor responses, between 13 and 20 (Fuortes and Hodgkin model) and between 20 and 48 (Baylor *et al.*'s model) steps were required — a seemingly implausible number. Payne and Howard's

model is simple and minimal in its complexity and assumptions: photoreceptor impulse responses are assumed to be logarithmically transformed Gaussians. The only variables are s , the sensitivity, which is fixed for a given cell, and μ and σ , the mean and standard deviation of the transformed Gaussian. Furthermore, the effect of light adaptation and temperature change can be modelled by changing a single parameter, μ , which changes the time scale of the response and could represent the rate constant of a single transduction process.

Mechanistically, as well as mathematically, the Payne and Howard model is a simple working hypothesis which fits available data and is strikingly similar to recent suggestions by Hamdorf and Kirschfeld⁴ arising from their experiments on fly photoreceptors. Payne and Howard assume that absorption of a photon by a rhodopsin molecule leads to release of a chemical transmitter, the concentration of which smoothly rises to a peak before declining. The transmitter molecules open the ion channels responsible for the change in membrane potential during illumination. In insects and those other invertebrates which have rhabdomeric photoreceptors, illumination causes depolarization. However, in order to adapt this model to vertebrate ciliary photoreceptors, which hyperpolarize during illumination, one need only assume that the transmitter molecules close, rather than open ion channels. The threshold transmitter concentration for activation varies

Peter G. Lillywhite is Lecturer in Zoology at Sheffield University.

from one ion channel to another according, it is assumed, to a gaussian distribution. This variability is responsible, according to the Payne and Howard model, for the long-lasting nature of the membrane potential produced by an instantaneous light pulse. The long latency of phototransduction, which can be tens of milliseconds for small responses in dark-adapted receptors, is assumed to be due to the time required for transmitter concentration to exceed the activation thresholds of the ion channels. Hamdorf and Kirschfeld recently pointed out that the rapidity of micro-spectrophotometrically observable pigment conversion (around 0.5 ms) and diffusion cannot account for the long latency of phototransduction in flies. In showing that simultaneous absorption of two photons by the same rhabdomere microvillus causes a discrete reduction in latency, Hamdorf and Kirschfeld concluded that each rhabdomere microvillus may represent a single functional unit, which generates transmitter molecules by means of a single enzymatically activated transduction step⁴. As in the Howard and Payne model, Hamdorf and Kirschfeld attribute the long latency to the time taken for the concentration of transmitter generated to exceed a threshold level.

An additional requirement for any realistic model of transduction in insects is that the first 3–5 ms of transduction latency consist of reversible events localized at the level of the photopigment molecule, since Hamdorf and Kirschfeld have recently shown that reconversion of metarhodopsin to rhodopsin within this period prevents the rhodopsin to metarhodopsin conversion having any effect⁵. Thus, no irreversible events, such as release of 'live' transmitter, occur in this period. It is attractive to assume that the remaining few milliseconds before onset of the response are indeed due to a single enzymatically activated generation of transmitter, as suggested by Hamdorf and Kirschfeld⁴ and as fitted by Payne and Howard's mathematical model. No transduction enzyme has yet been fully characterized, although there is some evidence that a specific protein is involved in the intermediate steps of phototransduction⁶.

It would be interesting to know how well the Payne and Howard model fits the responses of other invertebrate photoreceptors and whether it can be used to model the responses of vertebrate photoreceptors, because it predicts a much simpler model of transduction than the reaction chain hypotheses. □

Second sound in ³He

from P. V. E. McClintock

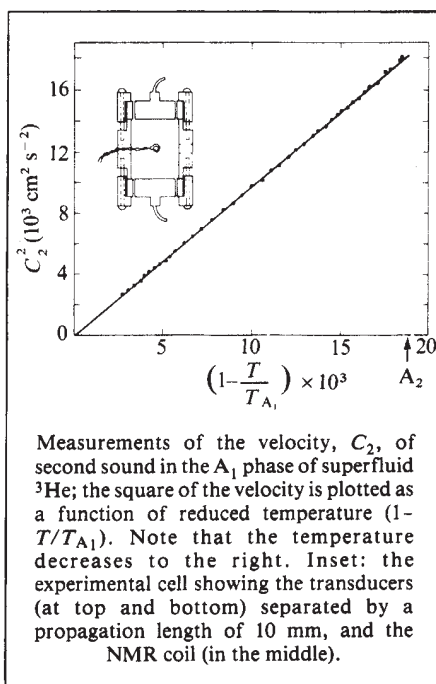
SECOND SOUND is an unusual type of propagating wave mode, which can occur in superfluids, involving fluctuations in the local temperature and entropy of the medium rather than in the local density and pressure as found in a conventional sound wave. The phenomenon has been investigated, both experimentally and theoretically, in the common helium isotope, ⁴He, over many years; but its observation in ³He, whose ultra-low temperature superfluidity is still a relatively recent (1972) discovery, was reported for the first time by Corruccini and Osheroff only last year (*Phys. Rev. Lett.* **45**, 2029).

At temperatures not far below their superfluid transitions (~2 K for ⁴He, 2 mK for ³He), both liquid heliums behave much as though they were composed of two separate but freely interpenetrating liquids: a superfluid component (anisotropic in the case of ³He) with no viscosity or entropy and a normal fluid component possessing the properties of a (fairly) normal liquid. This type of two-fluid system is able, in principle, to support two quite distinct wave modes. Where the two components vibrate in unison with each other, the result is an ordinary density wave known as 'first sound'. If, on the other hand, the two components move in antiphase, with the overall liquid density remaining essentially constant, the characteristic wave propagation velocity turns out to be quite different, corresponding to 'second sound'.

Although second sound is very readily generated and detected in superfluid ⁴He, calculations had suggested that the corresponding mode in the principal (A and B) phases of superfluid ³He would travel exceedingly slowly and be very strongly attenuated as it travelled through the liquid. Consequently, although second sound should, in principle, exist in ³He, it would quite probably prove impossible to detect it in practice. Liu predicted (*Phys. Rev. Lett.* **43**, 1740; 1979), however, that within a very narrow region of the phase diagram, the A₁ phase, the liquid might, in fact, be capable of supporting a freely propagating second sound wave — and it would be second sound of an entirely novel and particularly interesting kind.

The A₁ phase is formed under the influence of a strong magnetic field as an intermediate stage in the transition between normal liquid ³He and the superfluid A phase. The superfluid component of the latter can itself be considered as being made up of two sub-components in which the constituent pairs of ³He atoms lie with their nuclear magnetic moments lying either parallel or anti-parallel to an applied magnetic field. These 'up-spin' and 'down-

spin' superfluids seem, in many respects, to behave almost independently of each other. When normal liquid ³He is cooled in a magnetic field, the superfluid transition to the A phase is found to take place in two stages, at temperatures T_{A1} and T_{A2} whose separation is proportional to the applied field, and this is believed to correspond to the two superfluid sub-components having slightly different transition temperatures. Thus, for temperatures in the narrow range (typically 50 μ K) between T_{A1} and T_{A2} , only one of the spin species will be superfluid — this is known as the A₁ phase.



Liu pointed out that second sound in the A₁ phase, if it occurred, would carry an associated magnetic disturbance: counterflow of the normal and superfluid components would necessarily involve a spin density wave because only one preferentially aligned spin species would be present. He was able to predict that, for a hybrid second sound wave/spin density wave of this nature, the velocity would be much higher and the attenuation much smaller than would otherwise be anticipated for second sound in ³He.

It is this hybrid wave mode whose existence Corruccini and Osheroff have now successfully demonstrated at the Bell Telephone Laboratories, Murray Hill. They generated the mode by vibrating a capacitatively coupled porous membrane in the liquid. The precise mechanism by which this type of transducer operates is still somewhat ambiguous but, if one assumes that the membrane couples only to the normal fluid component, it is then clear that any motion of the membrane will result in a relative velocity of the normal

1. Fuortes, M.G.F. & Hodgkin, A.L. *J. Physiol., Lond.* **172**, 239 (1964).
2. Borsellino, A. & Fuortes, M.G.F., Smith, T.G. *Cold Spring Harb. Symp. quant. Biol.* **30**, 429 (1965).
3. Baylor, D.A. & Hodgkin, A.L., Lamb, T.D. *J. Physiol., Lond.* **242**, 685 (1974).
4. Hamdorf, K. & Kirschfeld, K. *Z. Naturforsch.* **35c**, 173 (1980).
5. Hamdorf, K. & Kirschfeld, K. *Nature* **283**, 859 (1980).
6. Pak, W.L. & Ostroy, S.E., Deland, M.C., Wu, C.-F. *Science* **194**, 956 (1976); Fung et al. *Proc. natn. Acad. Sci. U.S.A.* **78**, 152 (1981).

P. V. E. McClintock is in the Physics Department at the University of Lancaster, UK.

and superfluid components. The primary detection system depended on the same principles, but operating in reverse.

The received signals were weaker than anticipated, but yielded velocities which, as shown in the figure, appear to be in excellent agreement with the prediction, based on the theory of Liu, that the square of the velocity should be directly proportional to the reduced temperature ($1 - T/T_{A1}$). Some highly significant additional information was obtained from a small NMR coil placed in the middle of the experimental cell, which enabled the authors to monitor the magnetic disturbance which occurred as the second sound wave passed by on its way from the generator to the detector. On the basis of

careful analysis of the relative shapes of the NMR and acoustic signals they tentatively conclude that, quite contrary to expectation (Levin *Phys. Rev. Lett.* **34**, 1002; 1975), the paired spins forming the superfluid in the A_1 phase are in reality aligned 'up' along the magnetic field, rather than 'down'. This conclusion should be treated with caution until corroborated by some other means, but it is certainly intriguing.

Although several features of the experimental results still await detailed interpretation, there now seems no doubt that Liu's exotic wave mode exists in reality and that the work will lead to a significantly enhanced understanding of superfluid ^3He . $\square$

Ice core studies: dating the past to find the future

from W. Dansgaard

How did atmospheric carbon dioxide influence the climate in the past? Will the two large ice sheets, in Antarctica and Greenland, be stable enough to withstand a possible climatic warming? If not, how fast could they disintegrate? How old are they? How abrupt was the termination of the last interglacial, the Eem? When did it happen, and why? Did volcanic activity trigger that event? Is the Earth ripe for another catastrophic turn into glacial conditions? How did cosmic radiation change in the past? What were pre-industrial deposition rates of the various chemical elements that are now being emitted by modern industry?

These, and many other important problems may be partially, if not completely solved by investigation of the more than 3,000 m thick ice sheets which are proving richer sources of information about past atmospheric conditions than any other sequence of Quaternary sediments on Earth. All kinds of fall-out from the atmosphere, such as air-borne continental dust and biological material, volcanic debris, sea salts, cosmic particles and cosmic ray-produced isotopes, are deposited on the surface along with the snow and each year a new layer is added to the old. The snow pack is gradually compressed into solid ice with small cavities containing samples of atmospheric air. In the coldest areas of the ice sheets, the impurities remain in the ice as deep-frozen indicators of the chemical composition and the physical condition of the atmosphere at the time of deposition. Nothing is added, nothing runs off or is displaced, and no chemical reaction takes place; in fact, the composition of the ice layers changes only

by decay of the radioactive impurities and by extremely slow diffusion processes in the ice crystal lattice.

The ice layers sink into the ice sheet in an undisturbed sequence with continuous horizontal stretching and consequent thinning. In areas with no melting at the bedrock, the layers approach zero thickness close to the bottom. This is the reason why an ice core drilled through the coldest part of Greenland ice sheet makes it possible to establish continuous and detailed time series of many geophysical-chemical parameters, reaching several hundred thousand years back in time: CO_2 content in the atmosphere^{1,2}; climatic changes in terms of surface temperature³⁻⁵ and accumulation variations⁶; chemical composition of the atmosphere⁷; volcanic activity and its cooling effect in the troposphere⁸ and cosmic radiation flux¹. In central East Antarctica the oldest ice may even be an order of magnitude older than in Greenland, but the snow accumulation rate is so low⁹ — in some areas even zero or negative — that an Antarctic ice core does not necessarily represent a continuous sequence of deposition.

Dating of an ice core is prerequisite for proper interpretation of a data series. A number of dating methods have been developed on the basis of ice flow models, and analyses of isotopes, volcanic debris and continental dust in ice cores¹⁰. Most accurate are the stratigraphical methods, that is, detection and counting of the individual annual layers downwards from the surface. In areas with high accumulation ($> 25 \text{ cm ice yr}^{-1}$) and no melting, these methods allow absolute dating to within a few years per thousand if the ice core is in perfect physical condition. This

accuracy is required when high-frequency components of the time series are to be compared with similar features in other time series. For example, the acidity of annual layers could be reconciled with historical records of volcanic eruptions within the past millennium⁸ only because each layer was dated back to near the correct year of deposition¹⁰. So far, no ice core has been absolute-dated beyond AD 553.

Under favourable conditions the annual stratification will probably be detectable far beyond 10,000 yr BP, but, at present, ice from the early stages of the last glaciation must be dated by radioisotopes. However, the conventional low-level β -counting technique requires preparation of much larger amounts of ice than is available in an ice core¹¹. It is therefore a great step forwards that the limit of radioisotope detection has been lowered by orders of magnitude in recent years by means of an accelerator/mass spectrometer combination that identifies and detects the radioactive atoms themselves instead of the β -particles emitted when they decay¹²⁻¹⁵.

This makes it possible to perform ^{14}C dating back to 50,000 yr BP using as little as $1 \text{ cm}^3 \text{ CO}_2$ which is the approximate CO_2 content in 5 m of a 10 cm ice core. This core length may span only a few hundred annual layers at depths corresponding to 50,000 yr of age in the central parts of the large ice sheets¹⁶. Ice from the early stages of the last glaciation has to be dated by longer-lived isotopes, of which ^{36}Cl is the most promising. As the half life is relatively favourable (310,000 yr), a ± 3 per cent detection accuracy is within reach¹⁵; and an 0.5 m core increment is sufficient. Yet, the corresponding dating uncertainty will be as high as $\pm 13,000$ yr, even if the production rate of ^{36}Cl in the atmosphere is assumed to have remained constant. If the production rate at the time of deposition were 10 per cent different from the present value, it implies a dating error of no less than 43,000 yr.

This difficulty may be partly overcome if the ^{10}Be (half life 1.5×10^6 yr) concentration is measured in the same ice. The ^{36}Cl to ^{10}Be concentration ratio may then be used as an age indicator that is essentially independent of the production rate, because both isotopes are produced in the atmosphere by impact of cosmic radiation, and both of them are scavenged by precipitation after a short residence time in the atmosphere¹. Nevertheless, when accounting for another 3 per cent uncertainty on the ^{10}Be analysis, we end up with a total dating uncertainty of $\pm 30,000$ yr on an individual ice sample. Owing to the small amount of ice needed, it may be feasible to carry out independent datings on, say, 10 ice core samples of approximately the same age around 100,000 yr, thereby reducing the statistical uncertainty on the mean value to $\pm 10,000$ yr, which can be extremely useful for checking the validity of ice flow models. A complete $^{36}\text{Cl}/^{10}\text{Be}$ profile along the lower part of an

W. Dansgaard is in the Geophysical Isotope Laboratory of the University of Copenhagen.

ice core, in which the entire transition (120,000 to 60,000 yr BP?) from the last and really warm interglacial (the Eem) to full glacial conditions is revealed by oxygen isotope ratio changes, would probably lead to the long-needed climatic counterpart to the continental ice volume records derived from analyses of foraminifera in deep-sea cores¹⁷⁻¹⁹.

These records show that at the end of the Eem some 120,000 yr ago the amount of continental ice began to increase, and that it grew from less than today to more than double the present amount within less than 10,000 yr. The drastic climatic deterioration, which triggered off the enormous glacier advance, may have taken place within a much shorter time interval²⁰, perhaps a century or less, to judge from a Greenland isotope record²¹ and from pollen records²²⁻²⁴. However, these records are poorly dated before 50,000 BP, and the possibility of a discontinuity in the layer sequences must be taken into account, the more so as the suggested duration (one millenium only) of the cold period that succeeded the Eem is inconsistent with the build-up of huge amounts of continental ice.

It is extremely important to study why and how abruptly the Eemian interglacial came to an end, because it might change our conception of atmospheric stability, and describe a possible mode of termination of the present interglacial. The new radioisotope ice dating technique will not be able to reveal the duration of the past climatic shift, but a counting of annual layers, deposited in a continuous sequence throughout the shift, may be feasible by modified detection techniques, because such layers, each 1 mm thick, probably exist some 80 m above the bottom of the ice sheet in central Greenland¹⁶. □

The origin of cosmic gamma-ray bursts — new Soviet results

From B. J. Teegarden

SINCE their accidental discovery in 1973 by Vela spacecraft, gamma-ray bursts have continued to be a mysterious and intriguing phenomenon. Lasting from a few seconds to a few tens of seconds, these bright flashes of gamma radiation occur randomly every one to two weeks. With one possible exception the identification of the sources of these events remains unsolved.

A multi-spacecraft network including experiments designed in the United States, France and the Soviet Union has been used to time the arrival of the gamma-ray burst wavefront to determine the arrival direction of the radiation. The only burst for which a source identification was made is the very unusual event of 5 March 1979 which displayed a very sharp, intense radiation spike lasting only 0.1 second and was followed by 8 second oscillations lasting for approximately 100 seconds. The arrival direction of this burst coincided with the remnant of a supernova explosion (labelled N49) in the Large Magellanic Cloud, our nearest neighbouring galaxy. This, however, was a highly unusual event, and its identification is still not universally accepted. At this time no other gamma-ray burst source identifications have been made. This is in spite of the fact that arrival directions for a number of events have been precisely determined.

A number of different models have been put forth to explain the energy source of these bursts. Two of these may be briefly summarized: the first hypothesizes that a solid body falls onto the surface of a neutron star providing a nearly instantaneous release of energy. The second proposes that infalling material from either a companion star or the interstellar medium gradually builds up a surface layer at the poles of a neutron star. Sufficient accumulation of material over a period of time leads to an unstable condition wherein a thermonuclear explosion can occur. In either case the resultant heating of the atmosphere immediately above the surface of the neutron star is sufficient to cause the release of an intense burst of gamma radiation.

Some dramatic new observations of gamma-ray bursts have recently been reported by E. P. Mazets and his colleagues at the Ioffe Physical-Technical Institute in Leningrad (see this issue of *Nature*, p.378) that shed important new light on the question of the origin of these events. Mazets and his co-workers have measured

the spectra of a large number of bursts using detectors on board the Soviet Venera 11 and 12 spacecraft. They have found remarkable emission and absorption features in 26 of the 80-90 events that were measured. Most of the spectra display pronounced absorption in the 30-70 keV range. This has been interpreted as being due to a phenomenon known as cyclotron absorption.

It is expected that extremely strong magnetic fields will surround a neutron star. This follows if the magnetic flux is conserved during the collapse of the parent star down to the neutron star. A simple extrapolation from the parent star to its compact offspring yields a magnetic field in the 10^{12} gauss range, a value far beyond any achieved in the laboratory. An electron in such a field will be constrained to spiral tightly about the magnetic field line. Quantum theory predicts that the electrons will be confined to discrete energy levels and that the transitions between these levels will be accompanied by either the emission or absorption of radiation. In a 5×10^{12} gauss magnetic field the characteristic energy of the transition is ~ 50 keV. The absorption and emission features in gamma-ray burst spectra observed by Mazets and his colleagues are therefore consistent with cyclotron processes in the strong magnetic fields expected to exist near the surface of a neutron star.

In addition to the 'cyclotron' emission and absorption features, Mazets and co-workers, as well as a group at Goddard Space Flight Center in the US, have seen in a smaller group of events (about 6) a broad emission line in the 400-450 keV range. When a positron and electron encounter each other and annihilate, they give off a characteristic emission at 511 keV known as annihilation radiation. Any such radiation originating near the surface of a neutron star will lose energy as it escapes from this intense gravitational field. This energy change is known as a 'gravitational redshift'. Present models predict that the magnitude of this shift will be 15-20% for a neutron star whose mass is the same as the Sun. Any 511 keV annihilation radiation produced near a neutron star surface would, therefore, emerge from the star's gravitational field with an energy in the 400-450 keV range.

It should be emphasized that these results are very recent and that the interpretations discussed here are not necessarily unique. These Soviet results, however, seem to point very strongly towards a neutron star origin for cosmic gamma-ray bursts. □

B. J. Teegarden is in the Laboratory for High Energy Astrophysics NASA/Goddard Space Flight Center.

- Oeschger, H. *Proc. Symp. chem. and radioact. Dating* Am. Chem. Soc., Houston, Texas (March 1980).
- Delmas, R.J., Ascencio, J.-M. & Legrand, M. *Nature* **284**, 155 (1980).
- Johnsen, S.J., Dansgaard, W., Clausen, H.B. & Langway, C.C. *Nature* **235**, 429 (1972).
- Dansgaard, W. *et al.* *Nature* **255**, 24 (1975).
- Lorius, C., Merlivat, L., Jouzel, J. & Pourchet, M. *Nature* **280**, 644 (1979).
- Reeh, N. *et al.* *J. Glaciol.* **20**, 27 (1978).
- Heron, M.M., Langway, C.C., Weiss, H.V. & Cragin, J.H. *Geochim. cosmochim. Acta* **41**, 915 (1977).
- Hammer, C.U., Clausen, H.B. & Dansgaard, W. *Nature* **288**, 230 (1980).
- Barkov, N.I. *US Antarc. J.* **10**, No. 2, 55 (1975).
- Hammer, C.U. *et al.* *J. Glaciol.* **20**, 3 (1978).
- Oeschger, H., Stauffer, B., Brucher, P. & Moell, M. *J. Glaciol.* **17**, 117 (1976).
- Proc. 1st Conf. on Radiocarbon Dating with Accelerators* Univ. Rochester (ed. Gove, H.E.) (1978).
- Raisbeck, G.M., Yiou, F., Frumau, M., Lieuvain, M. & Loiseaux, J.M. *Nature* **275**, 731 (1978).
- Bennett, C.I. *et al.* *Science* **201**, 345 (1978).
- Finkel, R.C., Nishizumi, K., Elmore, D., Ferraro, R.D. & Gove, H.E. *Geophys. Res. Lett.* **7**, 983 (1980).
- Dansgaard, W., Johnsen, S.J., Clausen, H.B. & Gundestrup, N. *Meddelelser om Gronland* **197**, No. 2, 1 (1973).
- Emiliani, C. *J. Geol.* **74**, 109 (1966).
- Hays, J.D., Imbrie, J. & Shackleton, N.J. *Science* **194**, 1121 (1976).
- Duplessy, J.-C. in *Climatic Change* (ed. Gribbin, J.) 46 (Cambridge Univ. Press, 1978).
- Dansgaard, W. & Duplessy, J.-C. *Boreas* **10**, No. 2 (1981).
- Dansgaard, W., Johnsen, S.J., Clausen, H.B. & Langway, C.C. *Quat. Res.* **2**, 296 (1972).
- Andersen, S.T. *Medd. Dansk Geol. Foren.* **15**, 90 (1969).
- Wijmstra, T.A. *Acta bot. neerl.* **18**, 511 (1969).
- Woillard, G. *Quat. Res.* **9**, 1 (1978).

Plasmids and bacterial pathogens

from J.R. Saunders

THE way in which plasmids confer pathogenicity characteristics on host bacteria is of considerable concern in both developed and developing countries, and provided the central theme at a recent conference* in the Dominican Republic. Various aspects of plasmid biology were covered, ranging from the mechanisms of genetic transposition and plasmid replication to the epidemiology of antibiotic-resistant bacteria in human and animal populations.

Plasmids probably determine only a small proportion of all bacterial virulence traits. However, they may provide sufficient additional genetic information to convert relatively innocuous organisms into pathogens. This is particularly true of enteropathogenic strains of *Escherichia coli*, the prime cause of travellers' diarrhoea and other economically important diseases of humans and domesticated animals. The distressing symptoms of these conditions are largely produced by one or more plasmid-specified enterotoxins, which include a single type of heat-labile toxin (LT) and probably at least two distinct heat-stable toxins (ST). S. Falkow (Stanford) reported a striking similarity between the gene for LT and that for the cholera toxin, despite the fact that the former is always plasmid-borne in *E. coli* and the latter chromosomally determined in *Vibrio cholerae*. The difference in disease patterns, apparently specified by the same virulence determinant, may reflect the different methods by which the two causative agents actually deliver the toxin to their human hosts. This emphasizes how heavily the expression of pathogenicity determinants depends on the genetic and physiological infrastructure into which they become inserted. It is interesting in this context that certain characteristics, such as particular chromosomally determined somatic (O) antigen types, are found habitually among enterotoxigenic strains of *E. coli*.

Haemolytic strains of *E. coli* can frequently be isolated from infected body fluids. However, the exact role of haemolysins in the pathogenicity of these bacteria is not clear. In *E. coli*, α -haemolysins are determined either by the chromosome or by Hly plasmids. Such plasmids isolated from different strains have been shown by W. Goebel (Wurzburg) to share a common haemolysin determinant. This consists of three cistrons, *cisB* (1,500 base pairs) determining a protein located in the outer membrane and probably responsible for the transport of haemolysin, *cisA* (2,500 base pairs) encoding a precursor of α -haemolysin, and *cisC* (about 500 base pairs) which is apparently required for the

conversion of the *cisA* product to active α -haemolysin. Despite the wide distribution of the Hly determinant it does not seem to constitute part of a transposon (see De La Cruz *et al. J. Bact.* **143**, 825; 1980). This contrasts with one of the ST genes which is known to be transposable (So *et al. Nature* **277**, 453; 1979).

The ability of enteric bacteria to adhere to the gut wall is a primary requirement for their pathogenicity. The best studied of such plasmid-mediated determinants are those for the filamentous surface protein antigens K88 and K99, which enable enteropathogenic *E. coli* to adhere to the mucosa of the small intestine of pigs and calves respectively. Similar pilus-like structures, termed colonization factors (CFAI and II), are plasmid-encoded in human enteropathogenic *E. coli*. Two groups (J. van Embden, Bilthoven and G. Dougan, Dublin) reported the cloning and genetic analysis of the K88 and K99 antigens. Dougan has found that at least four genes, organized into two operons, are involved in the expression of the K88 antigen in *E. coli* minicells. Operon I comprises three genes transcribed from *adhA* (encoding a protein of molecular weight 70,000 which may be the basal structure for membrane attachment of the antigen) through *adhB* (29,000-MW protein) to *adhC* (17,000-MW protein). *adhC* may encode a positive regulator for operon II containing a single gene *adhD* determining the 23,000-MW subunit of the K88 antigen itself.

In addition to adhering to the intestine, some enteric pathogens are able to invade the gut tissues. D. Kopecko (Washington) described plasmids of about 120 megadaltons in *Shigella sonnei*. Loss of these plasmids results in transition of host strains from the virulent, smooth colonial form I to a virulent, rough colonial form II. Such plasmids apparently determine the synthesis of the form I somatic (O) antigen required for invasion of colonic epithelium.

The potential of bacteria as invasive pathogens is obviously enhanced by their ability to withstand the antibacterial properties of serum. A number of unrelated antibiotic-resistance plasmids and temperate bacteriophages, such as λ , confer on host bacteria resistance to complement (activated by either classical or alternative pathways). K. Timmis (Berlin) showed that serum resistance conferred by the conjugative plasmid R6-5 is determined by the product of the *traT* gene. This is a 25,000-MW protein responsible for the surface exclusion (prevention of conjugative mating with related donors) and present as 20,000 copies in the outer membrane of *E. coli* carrying R6-5. M. Binns (St Louis) has demonstrated that the *traT* product of another conjugative

plasmid, R100, impairs the formation of the terminal complement complex. Furthermore, *traT*-containing cells remain resistant to complement even if they have been pretreated with antibodies to *traT*. This suggests that the *traT* protein does not actively protect bacteria against serum but has a passive effect, possibly involving subtle alteration of the architecture of the cell envelope. In addition, studies by Timmis of mutants in the *traT* gene indicate that the serum-resistance and surface-exclusion properties of *traT* protein may be distinct. Whatever the mechanisms involved, serum-resistance determinants, such as *traT*, may provide a selective advantage leading to the maintenance of conjugative resistance plasmids in pathogenic bacteria.

An important prerequisite for invasive pathogens is the ability to proliferate in the low free-iron conditions prevailing in animal tissues. A significant proportion of *E. coli* strains causing bacteraemia in humans or animals are found to carry ColV plasmids. In addition to encoding colicin V production and serum resistance, such plasmids specify accessory iron-sequestering mechanisms. P. Williams (Leicester) has shown that they determine an extracellular iron-chelating compound of low molecular weight which competes successfully with transferrin for iron *in vivo*. In addition, they specify, together with chromosomally determined components, an active iron uptake system. The possession of these extra mechanisms for acquiring free iron is thus sufficient to allow growth of *E. coli* in the tissues of infected animals.

Knowledge gained from *E. coli* should facilitate the analysis of more complex, chromosomally determined pathogenic mechanisms in less tractable organisms. And cloning of virulence determinants provides the promise of effective vaccines against enteric and other infections.

J.R. Saunders is in the Department of Microbiology, University of Liverpool.



100 years ago

PHYSICAL NOTES

M. Plantamour continues to study with his sensitive levels the phenomena of periodic rise and fall of the ground which he has observed in Switzerland. He believes he has established a connection between these periods and those of the changes of temperature of the earth's surface, there being an annual change of level in an east-west direction corresponding with the mean temperatures of the surface during the year.

From *Nature* **23**, 31 March, 517, 1881.

*On the 'Molecular Biology, Pathogenicity and Ecology of Bacterial Plasmids' (organized by S.B. Levy, R.C. Clowes and E. Koenig), held from January 5th to 9th, 1981 in Santo Domingo, Dominican Republic.

Superluminal radio sources

from Guy Pooley

THERE was some understandable disbelief when it was reported about 10 years ago that some radio sources appeared to expand at velocities greater than that of light. Astronomers, though, are used to the impossible and since then there have been advances in both theory and observations. The article by Pearson *et al.* of the Cal Tech radio astronomy group (see this issue of *Nature*, p.365) describes observations of one particular quasar, 3C273, made with the US very long baseline interferometry (VLBI) network. They have applied the techniques of map-making developed by Readhead and Wilkinson at Cal Tech using so-called 'closure phase' measurements.

These observations all require the highest possible angular resolution — 1 milli-arcsecond or better — that at present can only be achieved by radio astronomers using transcontinental or inter-continental baselines. With these systems, it is not yet possible to measure absolute phase differences between signals arriving at the different stations; uncertainties in the clocks, geometry of the Earth, atmospheric paths and so on are too large. But if more than two receiving stations are available, there is some redundancy in the phase information, and the more stations in use the greater are the constraints that this places on the data. Pearson *et al.* used four, sometimes five stations and therefore six or ten baselines. Earlier long-baseline observers had to be content with fringe amplitude information only, and this was used to constrain models for the distribution of the radio sources on the sky; but it could not, for example, be used to distinguish between models which differed by a rotation of 180°.

The high apparent transverse velocities have been detected in several radio sources: quasars like 3C273, 3C279 and 3C345, and a galaxy, 3C120. These sources have been studied by VLBI techniques, until now without the use of closure phase. The results were perhaps most convincing in the case of 3C345, which could be modelled well by a simple double source and showed $v_{\text{apparent}} \approx 7c$.

3C273 is a source of some complexity. On the large scale, it has coincident radio and optical 'jets' some 20 seconds of arc (80 kiloparsecs) long. This jet is of course not detected in the long-baseline study, which shows the expanding feature some 0.01 seconds from the (presumed) central object. This feature is apparently moving with a transverse velocity of about 10c, rather more than had been estimated from the phaseless observations. The orientations of the large- and small-scale structures differ by about 20°.

One thing that cannot be done with closure phase is to find the absolute position of the source. It is therefore not yet possible to say with certainty which

features in the maps are 'stationary'.

The theoretical problems are not entirely new; Nova Persei (1901) gave similar difficulties, which were explained in terms of simultaneous excitation of separated clouds by a shock-front. Blandford, McKee and Rees reviewed the possibilities for the expansion of extragalactic radio sources (*Nature* **267**, 211; 1977). The range of possibilities is large. There are the 'acausal' models such as the 'Christmas tree' version involving independent sources whose unrelated eruptions mimic expansion and there are models depending on the variation of propagation conditions either close to the source or in the intervening space (gravitational lenses, or scintillation in some intervening medium of galaxy). Neither of these general types seems to explain naturally the common features, particularly the fact that the sources always seem to expand.

The most popular models involve ejection of matter in some form of relativistic jet; there are several variants in which this can lead to high apparent transverse velocities. One particular attraction lies in the fact that the brightness temperatures of some compact radio sources, as implied by their rapid variations and inferred small sizes, are already uncomfortably high for the incoherent synchrotron process; a relativistic expansion eases this problem.

Scheuer and Readhead (*Nature* **277**, 182; 1979) have developed one attractive version of this model. In it, they postulate that radio emitting objects are emitted repeatedly along one axis from the central object. Most of the light comes from the central, stationary object. The relativistic γ for the approaching component must range up to around 10, and any receding components would be practically undetectable. In their model, it is those quasars or galaxies whose axes are aligned almost along the line of sight for which dramatic effects are observed; those which are not so aligned are radio quiet, or at least radio dim, and may form central components of 'classical' double sources.

Pearson *et al.* make the point that the Scheuer-Readhead model is in some danger because of the dearth of radio-quiet quasars of comparable magnitude to that of 3C273; this is one of the questions for which careful statistics need assembling, and this can be done only by systematic mapping of expanding sources of this type. Further questions can now be asked; is the expansion uniform? When the size doubles, it might be surprising if the speed does remain constant. What are the statistics of a complete sample of quasars? Are the detailed structures consistent with the expanding jet models? The evidence is that rapid expansion is a very common feature among compact radio sources. □

Guy Pooley is in the Mullard Radio Astronomy Observatory, University of Cambridge.

DNA methylation and control of gene expression

from Tomas Lindahl

ABOUT 3% of cytosine residues in mammalian DNA are present as 5-methylcytosine, methylation of cytosine usually occurring soon after DNA synthesis. The existence of such signal groups has often stimulated speculation, but hard facts on their role have been scarce. However, recent experiments suggest a role for DNA methylation in regulating or blocking transcription of DNA. Such data further indicate that specific demethylation events may be necessary for gene expression during development and differentiation. A meeting on the subject was recently held at the University of Cologne* and one point that emerged was that an understanding of the role of DNA methylation in mammalian cells is unlikely to be obtained from studies on lower organisms and other model systems.

The relationship between DNA methylation and ribosomal gene function in several different organisms was reviewed

by A. Bird (MRC Mammalian Genome Unit, Edinburgh). Mouse cells carry about 100 ribosomal RNA genes per cell, occurring in two forms, unmethylated and heavily methylated. In various inbred mouse strains, between 5 and 30% of the sequences are methylated. Only the unmethylated ribosomal genes in mouse liver appear to be expressed, and the data point to an inverse correlation between DNA methylation and transcription. However, in *Xenopus laevis* all the chromosomal ribosomal DNA is heavily methylated, and this vertebrate clearly transcribes methylated sequences, although transcription may depend on the presence of an unmethylated region within the non-transcribed spacer between the ribosomal DNA genes. In contrast, *Drosophila* cells contain virtually undetectable amounts of 5-methylcytosine, so this organism is not likely to utilize DNA methylation for control of gene activity.

Many plant cells contain exceptionally high concentrations (up to 25%) of 5-methylcytosine in their DNA, and transcription of such methylated DNA does take

*A meeting on 'DNA Methylation and Genome Organization' (4-7 March, 1981) was organized by the Institute of Genetics, University of Cologne.

place (B. Deumling, Heidelberg). The reason for the heavy DNA methylation in plants remains unknown, but replacement of cytosine with a more radiation-resistant derivative could provide the cells with an increased tolerance to sunlight.

In animal cells, methylated cytosines are often present in relatively large amounts in non-transcribed satellite sequences, and fluctuation of the number of such regions in different cell types has often obscured more interesting variations in DNA methylation. This problem has been circumvented by examining the methylation of specific DNA sites within individual genes using restriction enzyme analysis and cloned DNA fragments as probes. Some of the most interesting recent data have come from work with DNA tumour viruses. L. Vardimon and W. Doerfler (Institute of Genetics, University of Cologne) showed for three different adenovirus type 2 transformed hamster cell lines, which contained the entire gene for the virus-coded 72K DNA binding protein, that the mRNA for the protein (as well as the protein itself) was present in only one of the cell lines. In the latter, the integrated viral DNA was unmethylated, as judged by its sensitivity to the *HpaII* restriction endonuclease, while the corresponding DNA sites in the two other cell lines were methylated. Similar results were obtained with a cloned viral DNA fragment methylated *in vitro* at a few cytosine residues. Furthermore, *in vitro* transcription studies (performed together with J. Corden, Strasbourg) indicated that the initiation of transcription of the gene for the 72K protein was severely inhibited by methylation of the DNA.

An oncogenic herpesvirus, Herpes saimiri, is similar to the adenoviruses in that the DNA in virus particles is completely unmethylated, while the largely non-expressed viral DNA present in virus-induced tumours or in transformed cell lines is heavily methylated (B. Fleckenstein and C. Kaschka-Dierich, Erlangen and R.C. Desrosiers, New England and Primate Research Center). As with the similar human lymphotropic herpesvirus, Epstein-Barr virus, the viral DNA in H.saimiri transformed cells is largely carried as multiple copies of nonintegrated plasmid DNA; S-adenosylhomocysteine treatment (which inhibits DNA methylation) markedly decreased the amount of viral DNA carried per cell.

As might be expected, there are exceptions to the rules. SV40 DNA from virions is apparently completely unmethylated. Microinjection experiments with SV40 DNA methylated *in vitro* with a highly purified rat liver methylase (which acted on all CpG sequences as well as several other C residues so that 2.5% of the cytosines were converted to 5-methylcytosine) seemed at variance with some of the data obtained for other tumour viruses (D. Simon and A. Graessmann, Berlin). On microinjection of 2 to 4 viral DNA molecules per host cell, there was apparently no

active demethylation of the SV40 DNA, but nevertheless, the methylated DNA induced SV40 determined antigens as effectively as nonmethylated DNA. Similar data were obtained with both permissive and nonpermissive host cells, so these experiments appear to suggest a difference in the sensitivity to DNA methylation between SV40 and larger DNA tumour viruses. The problems with this approach are that the incoming methylated DNA may have been replicated into nonmethylated daughter molecules, or the methylated DNA preparation may have contained a minority of unmethylated molecules which were transcribed.

An elegant system for the investigation of the effect of DNA methylation is provided by studying transformation of thymidine kinase deficient cells using cloned thymidine kinase genes. M. Wigler (Cold Spring Harbor Laboratory) used genes from Herpes simplex type 1 and from chicken, and compared the biological activity of unmethylated DNA and DNA methylated *in vitro* with *HpaII* methylase. The methylated genes showed distinctly lower transformation efficiency than the unmethylated controls, although the two kinds of molecules had the same frequency of incorporation. Methylation either in the promoter region or in the structural gene caused equal inhibition compared to unmethylated controls, so specific methylation of the promoter region did not seem to be of unique importance.

In the above experiments, the methylated cytosine residues of the incoming DNA, which were present in CpG sequences, appeared to serve as their own genetic determinants and were largely preserved for many generations. It seems clear that mammalian cells contain a maintenance methylase to retain the general methylation pattern of DNA. Similar data were presented by H. Cedar (Hebrew University, Jerusalem), who replaced dCTP with 5-methyl-dCTP in *in vitro* replication to obtain a very highly methylated DNA molecule. So far, only model experiments had been performed with the replicative form of Φ X174 DNA, containing one strand with all cytosines replaced by 5-methylcytosine. After introduction into hamster cells, where the DNA integrated at low efficiency, methylated sites at all CpG sequences seemed to be preserved while other 5-methylcytosines were replaced by unmodified cytosine on replication. This indicates that the maintenance methylase requires the CpG sequences for activity. However, the slow progress made in the characterization of mammalian DNA methylase(s) makes such observations seem somewhat preliminary.

Detailed restriction enzyme analysis of methylation sites within the human globin gene clusters (R.A. Flavell, National Institute for Medical Research, London) shows that many methylated residues persist also in genes that are being expressed, but globin DNA that is transcribed is

undermethylated compared to that present in sperm. These observations again seem in accord with the general idea that DNA methylation may play a role in controlling and preventing transcription. One way of overcoming inherent problems that arise from using DNA cloned in *Escherichia coli* in biological assays such as *in vitro* transcription is to use DNA heavily methylated *in vitro*, on the assumption that it is easier for a cell to specifically demethylate DNA than to methylate it, in analogy with the demethylation that apparently occurs with germ-line globin DNA on activation. It is not known how this demethylation is achieved, although several models are possible. For example, DNA may be replicated in a cellular compartment in which methylation at some sites does not occur, or specific DNA binding proteins may prevent the methylation at certain loci during replication. Alternatively, some 5-methylcytosine could be actively replaced by unmodified cytosine in the same way as misincorporated uracil residues are efficiently removed from DNA.

The role of DNA methylation in simpler organisms than mammalian cells is also being studied. In an important advance, R. Sager (Sidney Farber Cancer Institute, Boston) has shown that the maternal inheritance of chloroplast genes in *Chlamydomonas* results from the methylation of chloroplast DNA in female (but not in male) gametes, leading to the preferential degradation of the chloroplast DNA of male origin in zygotes. In cells carrying a *mat-1* mutation, the male chloroplast DNA also becomes methylated, and consequently the inheritance pattern is converted from maternal to biparental. M. Meselson (Harvard University) discussed the role of DNA adenine methylation for the direction of mismatch repair of phage λ heteroduplex molecules. Some, but not all, of various constructed mismatches having one strand methylated *in vitro* with a DNA methylase and the other strand unmethylated were susceptible to repair. While DNA sequence data on the regions around the mismatched residues are not yet available, it seems clear that the mismatches that are being repaired have to be present in a heteroduplex structure in which only one strand of the DNA is methylated, and that repair then occurs selectively in the unmethylated strand. Thus, neither fully methylated DNA nor unmethylated DNA is susceptible to mismatch repair, and it presumably only takes place in newly replicated DNA with the parental strand methylated. While mammalian cells probably also have the ability to perform mismatch repair, they do not seem to contain 6-methyladenine in their DNA, so it is an open question whether DNA methylation can guide mismatch repair events in higher organisms. □

Tomas Lindahl is Professor of Medical Biochemistry at the University of Gothenburg, Sweden, presently at the Imperial Cancer Research Fund Laboratories, London.

ARTICLES

Superluminal expansion of quasar 3C273

T. J. Pearson, S. C. Unwin, M. H. Cohen, R. P. Linfield, A. C. S. Readhead,
G. A. Seielstad, R. S. Simon & R. C. Walker*

Owens Valley Radio Observatory, California Institute of Technology, Pasadena, California 91125, USA

Maps of the radio structure of 3C273 show directly that it expanded with an apparent velocity 10 times the speed of light from mid-1977 to at least mid-1980.

THE radioastronomical technique of very long baseline interferometry (VLBI) achieves an angular resolution better than 1 mas, equivalent to a physical resolution of the order of light-years in even the most distant galaxies and quasars. At this resolution, changes in the structures of such objects can be easily detected. In four extreme cases, three quasars (3C273, 3C279 and 3C345) and one galaxy (3C120) have been seen to expand with an apparent speed more than twice the speed of light, c (ref. 1). The evidence for such 'superluminal' expansion has been based on rather limited data, and the interpretation has been uncertain; but the recent development of the VLBI technique of hybrid mapping² allows us to map the radio brightness distribution to show the expansion directly and unambiguously.

Table 1 Observations

Epoch	Antennas*	Frequency† (MHz)	Flux density‡ (Jy)
1977.56	KGFO	10,651	40.9
1978.24	KGFO	10,651	44.5
1978.92	KGFO	10,651	34.9
1979.44	BKGFO	10,651	30.7
1980.27	BKG O	10,651	27.0
1980.52	BKGFO	10,651	27.8
1977.92	GFOH	5,011	46.2
1978.51	GFOH	5,011	45.6
1979.25	B G OH	4,997	40.0
1979.92	BKGFO	5,009	36.0

* B, 100-m, Max-Planck-Institut für Radioastronomie, Effelsberg, FRG; K, 36-m, Haystack Observatory, Massachusetts; G, 43-m, National Radio Astronomy Observatory, Green Bank, West Virginia; F, 26-m, Harvard Radio Astronomy Station, Fort Davis, Texas; O, 40-m, Owens Valley Radio Observatory, Big Pine, California; H, 26-m, Hat Creek Radio Astronomy Observatory, Cassel, California.

† The centre frequency of the 2-MHz band.

‡ Measured during the observation; typical uncertainty 0.5 Jy.

The 10.65-GHz observations were made in left-circularly polarized radiation, and the 5.0-GHz observations were made in linearly polarized radiation with the E-vector in PA 90°.

This article presents maps of the quasar 3C273 which show that it was expanding with an apparent velocity $v = (9.6 \pm 0.8)c$ between 1977.5 and 1980.5. (We assume $H_0 = 55 \text{ km s}^{-1} \text{ Mpc}^{-1}$ and $q_0 = 0.05$ throughout.)

3C273 is a 13-mag quasar with a redshift of 0.158 (ref. 3). An optical jet extends from a distance of ~ 10 arcs to ~ 25 arcs from the quasar in position angle (PA) -137° (ref. 4). The projected length of the jet is 84 kpc. The radio emission comes principally from two locations: a compact component (3C273B) coincident with the quasar, and a more diffuse component

(3C273A) near the end of the jet⁵⁻⁷. The compact component 3C273B consists of a flat-spectrum core with a steeper-spectrum extension in PA -117° (ref. 8). This direction is approximately that of the superluminal expansion. As the angular scale is increased, the predominant position angle rotates until it is more closely aligned with the optical jet⁹. Recent observations of 3C273 in all wavebands have been reviewed by Ulrich¹⁰.

3C273B was one of the first sources to be studied with VLBI^{11,12} and has been the subject of many subsequent observations. Changes in the visibility function were first observed in 1971^{13,14}. Although these observations were confined to a single baseline, it was immediately clear that the changes could be explained by a simple expansion. The data permitted other explanations, however, including models in which the centres of emission remained fixed while varying in intensity¹⁵. With more complete visibility data¹⁶⁻¹⁸, it became clear that the structure was more complex than a simple linear model of two or three components, and that the overall extent of the brightness distribution continued to increase. To avoid some of the ambiguities inherent in model-fitting, Cohen *et al.*¹ examined the visibility function measured between 1970 and 1977, and found that the first minimum (that at shortest projected baseline) in the visibility amplitude moved steadily towards smaller baselines, indicating a steady increase in the size of the source. They took as a measure of the angular size the quantity θ^* , defined as $1/(2w)$, where w is the projected baseline (in wavelengths) at the first minimum. If the source were a simple double, θ^* would be the separation of the two components. A later analysis¹⁹ of all the data available up to 1977.41 showed that the time-dependence of θ^* approximated a straight line with a slope of $0.41 \pm 0.04 \text{ mas yr}^{-1}$, corresponding to an apparent linear expansion rate of $v/c = 5.2 \pm 0.5$. This steady increase in θ^* (Fig. 16 of ref. 19) has provided the strongest evidence for superluminal expansion in 3C273. We know that 3C273 was not accurately represented by a simple double model during the period spanned by these observations, so θ^* was not a true measure of its angular size. We should not be surprised therefore that the maps presented here show that the true expansion speed differs from that inferred from the measurements of θ^* .

Maps

Since 1977 we have been monitoring 3C273 at two frequencies, 10.65 GHz and 5.0 GHz, using four or five antennas. The new observations are better than the old for several reasons: (1) we have used more antennas, sampling the visibility function at more spatial frequencies; (2) we have when possible included an antenna in Europe in addition to American antennas, thereby improving the resolution; (3) we have measured closure phase²⁰ in addition to visibility amplitude. These improvements enable us to make hybrid maps of the brightness distribution². The increase in the quantity of data has eliminated the ambiguities of model-fitting. The first maps at these two frequencies have already been published^{8,21}, and we have now observed a total of six epochs at 10.65 GHz and four epochs at 5.0 GHz (Table 1).

* Present address: National Radio Astronomy Observatory, Edgemont Road, Charlottesville, Virginia 22901, USA.

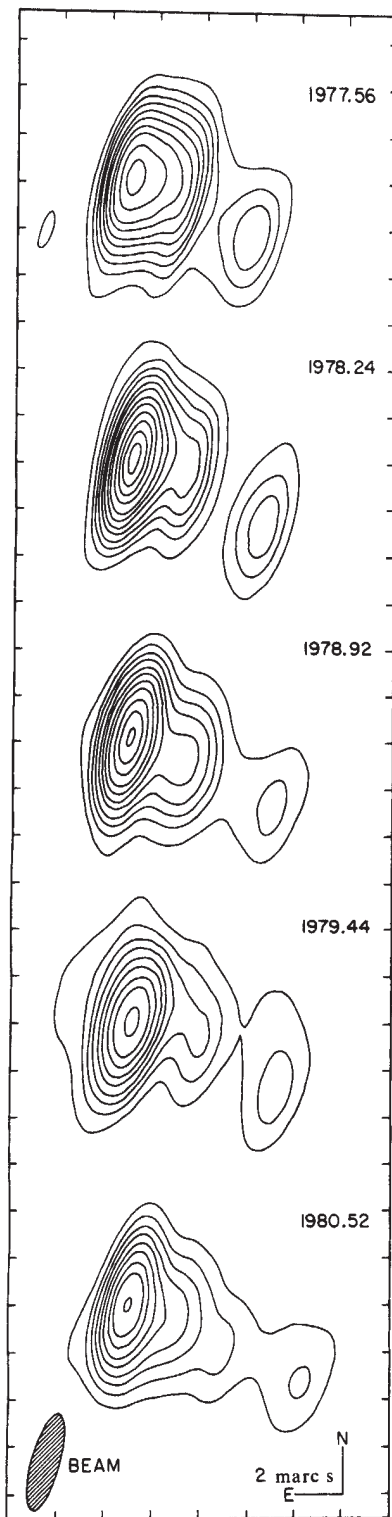


Fig. 1 Maps of 3C273B at 10.65 GHz at five epochs (see Table 1). The same restoring beam (an elliptical gaussian with FWHM 4.2×1.2 marc s, with the major axis in PA -15° , indicated by the hatched ellipse) is used for all the maps. The dynamic range of the maps is about 20:1. The contour levels for all the maps are 1.2, 2.4, 3.6, 5, 7, 9, 12, 15, 20, 25, 30 and 35×10^9 K. The absolute position of the source cannot be accurately determined from the data; here the maps have been lined up on the sharp eastern edge to show the motion of the weaker western component. The map for epoch 1977.56 is from ref. 8, which gives more details of the map-making procedure.

Full details of these observations will be published elsewhere; here we describe only the 10.65-GHz maps and the superluminal expansion that they reveal.

Figure 1 shows maps made from five of the six 10.65-GHz observations. It was not possible to make a reliable map from the observation of 1980.27 owing to a failure at one of the antennas. The beam (hatched ellipse in Fig. 1) is very elongated because 3C273 is close to the celestial equator, where the presently available network of antennas has very poor north-south resolution. Fortunately the principal extension of the source is close to the direction of maximum resolution. Two of the experiments included an antenna in Europe as well as those in America (see Table 1). For these, the east-west resolution was improved by a factor of two. To make the maps directly comparable, however, the restoring beam appropriate to the lowest-resolution observations was used for all the maps.

The five maps in Fig. 1 span a period of 3 yr. All the maps show a bright eastern peak and a lower-brightness extension to the west. In all cases there is a local maximum in the western

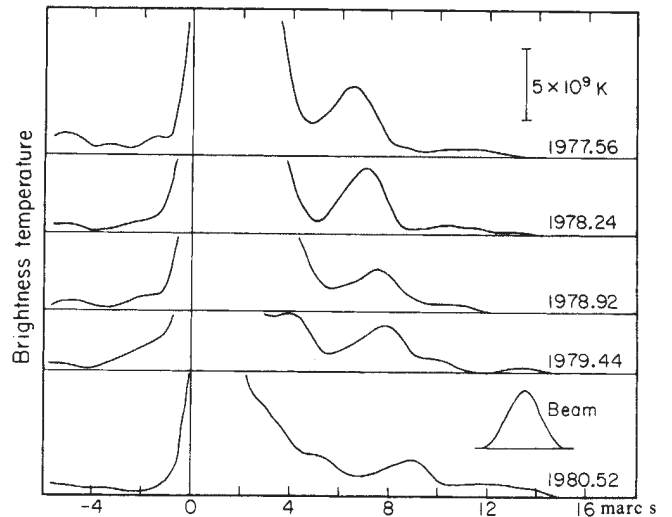


Fig. 2 Brightness-temperature profiles through the maps of Fig. 1 along the principal axis of the source (PA -116°). The profiles have been truncated to show the evolution of the weak western component, and they have been aligned at the half-power point on the eastern edge, labelled 0 on the abscissa. The vertical displacement between the profiles is proportional to the difference in epoch. The profile of the restoring beam (FWHM 1.42 marc s) is indicated.

extension between 6 and 8 marc s from the main peak. The most striking difference between the maps is the steady increase in the separation of this local maximum from the main peak. The position angle defined by the main peak and the secondary maximum has not changed significantly; in all the maps it is $-116^\circ \pm 2^\circ$. The maps also show that the source is not straight but curves to the south with increasing separation from the main peak. This phenomenon has been noted previously^{8,9,18}.

Expansion velocity

The profiles in Fig. 2 show that the separation of the western maximum from the main peak increased from ~ 6 marc s to ~ 8 marc s between 1977.5 and 1980.5. The time-dependence of the separation, θ , is shown in Fig. 3. In addition to the 10.65-GHz points, one 5-GHz measurement is included to show that the size of the source is the same at both frequencies. (The other 5-GHz observations do not have sufficient resolution to be directly comparable, but they are consistent with this conclusion.) θ was determined by measuring the separation of the peak of the western maximum from the half-power point in the sharp eastern edge, and then subtracting half the FWHM of the restoring beam (0.71 marc s) to correct for resolution. This measure is less affected by changes in the internal structure of the eastern component than a separation measured to the peak

of the eastern component. In all cases the measurement was made along a line with PA -116° .

The values of θ all fall remarkably close to a straight line in Fig. 3, indicating that the expansion is linear, with no evidence for acceleration or deceleration. The angular expansion rate, derived by a least-squares fit of a straight line (omitting the 5.0-GHz point), is 0.76 ± 0.04 marc s yr $^{-1}$, corresponding to a linear expansion rate of $v/c = 9.6 \pm 0.5$. The extrapolated intercept $\theta = 0$ is at 1970.1 ± 0.6 . If the expansion speed has not changed, one might expect to see an increase in total flux density at this epoch, corresponding to the formation of the new component. No such increase was seen²². The flux density of 3C273 has irregular variations with a time scale of one or two years, but it is not possible to correlate changes in the flux density with structural changes.

θ is larger than the values of θ^* extrapolated from the older data, but this is due to the difference in measuring technique. θ^* , being determined directly from the visibility function, is not related in a simple way to the brightness distribution except when the source is a simple double. If the expansion were conformal (if the brightness distribution retained the same shape but simply increased in scale), θ would differ from θ^* by a constant factor, but the real behaviour is more complicated.

Acceleration?

The angular size of 3C273B derived by extrapolating the expansion rate shown in Fig. 3 back to the early 1970s is smaller than the size then observed. This implies that either the expansion speed has increased, or the expansion has not been conformal. It is possible to devise non-conformal models that can explain the qualitative features of the observations without requiring an acceleration. In triple models, unlike doubles, the positions of the minima depend not only on the separations of the components but also on their relative intensities. One such model involves one fixed component and two components both moving with the same velocity. Suppose that at the earlier epochs (1970–75) the source consisted of three principal compact components: a core C at the eastern end, with flux density S_C , and two components A and B moving away from the core at $10c$ in PA -116° , with flux densities S_A and S_B . Let A be the more distant component. When the separation of A from C is much greater than the separation of B from C, the first visibility minimum measures the distance between A and the centroid of B and C. The expansion rate derived from θ^* is then $10cS_C/(S_B + S_C)$. This rate was $5.6c$ (ref. 19), implying that $S_B \approx S_C$ in 1970–75. Component A has since decayed below the level of detectability, and the new maps measure the separation of B from C which is increasing at $10c$. Of course, this model is an oversimplification, as the maps show additional components lying between B and C.

It is not possible to eliminate models involving acceleration, but a constant-velocity model seems to be favoured by the constancy of the expansion rate shown by the maps of Fig. 1. If the constant-velocity model is correct, we would expect further components to appear in maps made during the next few years.

Relativistic jet

Although several explanations have been suggested for superluminal expansion^{23,24} we shall discuss only the model²⁵ in which the expansion is presumed to represent a physical motion almost directly towards the observer; we are looking down a jet of relativistic material expelled from the quasar. The eastern end of 3C273B is the 'core', the point at which the jet becomes optically thick. In a steady flow, this point remains fixed even though the radiating material has a high velocity. The moving (western) component is a 'knot' propagating outwards along the jet. At large distances, the jet becomes visible as the optical jet, and the region of impact on the external medium is represented by the radio source 3C273A. This model accounts for (1) the milli-arcsecond morphology of 3C273B; (2) the fact that the expansion is roughly along the source axis; and (3) the approximate alignment between the axis of 3C273B and the optical jet.

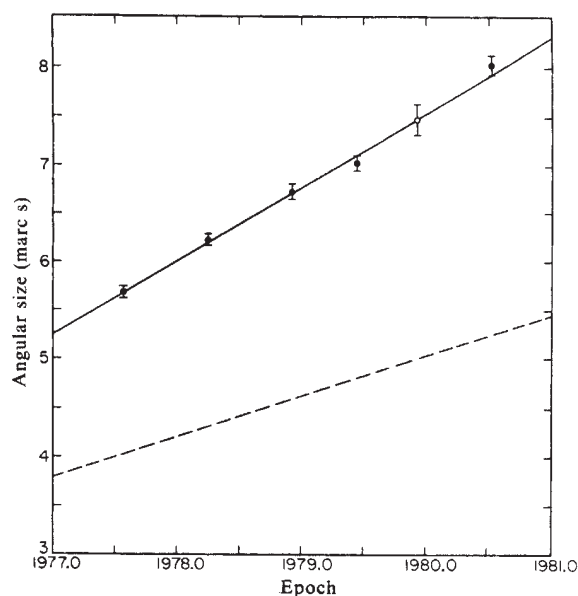


Fig. 3 Angular size of 3C273 as a function of time, measured as described in the text. Five 10.65-GHz measurements (●) and one 5.0-GHz measurement from 1979.92 (○) are shown. The dashed line at the bottom is an extrapolation of the expansion rate derived from observations of θ^* before 1977 (ref. 19). The solid line is a least-squares fit to the 10.65-GHz measurements, and has a slope of 0.76 marc s yr $^{-1}$.

(The difference of 20° between the radio axis and the optical axis is attributed to curvature in the jet⁹.) If the constant-velocity interpretation of the older data is correct, then at least two knots have been ejected from the core in a period of ~ 10 yr. Alternatively, an accelerating expansion can be accommodated in this model: it could be due to real acceleration in the flow along the jet or to a curvature of the jet.

To account for an apparent transverse velocity v as large as $9.6c$, the angle ϕ between the jet and the line of sight must be small. The apparent velocity v is related to ϕ and to v_0 , the space velocity of the moving feature, by

$$v = v_0 \sin \phi / \{1 - (v_0/c) \cos \phi\}$$

(ref. 26). If $v = 9.6c$, this equation requires that $\phi < 11.9^\circ$ ($\cos \phi > (v^2 - c^2)/(v^2 + c^2)$) if $v_0 < c$, and v_0 is minimized when $\phi = 6.0^\circ$. The probability that ϕ should be as small as 11.9° in a randomly selected quasar is 1%, or 2% if the jet is two-sided.

3C273 is far from being a randomly chosen quasar: it is the closest quasar in the 3CR sample²⁷ and has an intrinsic optical intensity ~ 4 times the intensity of the other quasars in the sample. It is surprising that it should also be exceptional in having a jet pointing almost directly at Earth. A possible explanation is that the optical radiation, like the radio, is not isotropic. There is no direct evidence for this, but beaming of the optical continuum cannot be ruled out. If the optical radiation is isotropic, then the small angle between the radio jet and the line of sight causes grave difficulties for the simple relativistic theories. For example, Scheuer and Readhead²⁸ attribute the difference in radio intensity between optically selected and radio-selected quasars to the relativistic beaming of the radio radiation—radio-selected quasars point towards Earth. We would expect on this model to see 50–100 radio-quiet quasars of comparable optical magnitude to 3C273, which we do not. It may be necessary to consider other models in which the radio radiation is not confined to so small an angle; one such model has been suggested by Rees²⁹.

Conclusion

The maps presented in Fig. 1 provide the first direct and unambiguous evidence of superluminal expansion in any radio source. For the period of 3 yr spanned by the observations the superluminal expansion could be attributed to the movement of

a single knot away from the nucleus along the jet. The apparent constant velocity of 9.6c observed in this period is an important constraint on theories of apparent superluminal expansion. Further observations of both 3C273 and other superluminal sources are needed to show whether the expansion is always at a constant velocity; the present results show that such analyses must be based on maps, rather than on the motion of features in the visibility function.

The observations described here were made under the auspices of the United States VLBI Network Users Group. We

Received 3 February; accepted 25 February 1981.

1. Cohen, M. H. *et al. Nature* **268**, 405–409 (1977).
2. Readhead, A. C. S. & Wilkinson, P. N. *Astrophys. J.* **223**, 25–36 (1978).
3. Schmidt, M. *Nature* **197**, 1040 (1963).
4. Greenstein, J. L. & Schmidt, M. *Astrophys. J.* **140**, 1–34 (1964).
5. Hazard, C., Mackey, M. B. & Shimmins, A. J. *Nature* **197**, 1037–1039 (1963).
6. Conway, R. G. & Stannard, D. *Nature* **255**, 310–312 (1975).
7. Perley, R. A., Fomalont, E. B. & Johnston, K. J. *Astr. J.* **85**, 649–658 (1980).
8. Readhead, A. C. S. *et al. Astrophys. J.* **231**, 299–306 (1979).
9. Readhead, A. C. S. *et al. Nature* **276**, 768–771 (1978).
10. Ulrich, M.-H. *Space Sci. Rev.* (in the press).
11. Broten, N. W. *et al. Nature* **215**, 38 (1967).
12. Clark, B. G., Cohen, M. H. & Jauncey, D. L. *Astrophys. J. Lett.* **149**, L151–L152 (1967).
13. Whitney, A. R. *et al. Science* **173**, 225–229 (1971).
14. Cohen, M. H. *et al. Astrophys. J.* **170**, 207–217 (1971).

thank the many people involved in scheduling the observations, recording the data at the participating observatories, and processing the tapes, for indispensable contributions to this paper. The work was supported by the NSF via grants AST 77-00247 and AST 79-13249 to the Owens Valley Radio Observatory; the National Radio Astronomy Observatory is operated by Associated Universities, Inc., under contract with the NSF; the astronomical VLBI programs at Harvard Radio Astronomy Station, Hat Creek Radio Astronomy Observatory and Haystack Observatory are supported by the NSF.

15. Kellermann, K. I. *et al. Astrophys. J. Lett.* **189**, L19–L22 (1974).
16. Legg, T. H. *et al. Astrophys. J.* **211**, 21–30 (1977).
17. Schilizzi, R. T. *et al. Astrophys. J.* **201**, 263–274 (1975).
18. Niell, A. E., Kellermann, K. I., Clark, B. G. & Shaffer, D. B. *Astrophys. J. Lett.* **197**, L109–L112 (1975).
19. Seielstad, G. A. *et al. Astrophys. J.* **229**, 53–72 (1979).
20. Rogers, A. E. E. *et al. Astrophys. J.* **193**, 293–301 (1974).
21. Cohen, M. H. *et al. Astrophys. J.* **231**, 293–298 (1979).
22. Andrew, B. H., MacLeod, J. M., Harvey, G. A. & Medd, W. J. *Astr. J.* **83**, 863–899 (1978).
23. Blandford, R. D., McKee, C. F. & Rees, M. J. *Nature* **267**, 211–216 (1977).
24. Marscher, A. P. & Scott, J. S. *Publ. astr. Soc. Pacif.* **92**, 127–133 (1980).
25. Blandford, R. D. & Königl, A. *Astrophys. J.* **232**, 34–48 (1979).
26. Ginzburg, V. L. & Syrovatskii, S. I. *A. Rev. Astr. Astrophys.* **7**, 375–420 (1969).
27. Schmidt, M. *Astrophys. J.* **151**, 393–409 (1968).
28. Scheuer, P. A. G. & Readhead, A. C. S. *Nature* **277**, 182–185 (1979).
29. Rees, M. J. *Proc. IAU Symp.* No. 94 (in the press).

Human immunoglobulin κ light-chain genes are deleted or rearranged in λ -producing B cells

Philip A. Hieter*, Stanley J. Korsmeyer†, Thomas A. Waldmann† & Philip Leder*

* Laboratory of Molecular Genetics, National Institute of Child Health and Human Development,
† Metabolism Branch, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20205, USA

We have examined 10 human λ light-chain-producing B lymphocytes and found that genes for the κ constant region have been deleted in each. Only one of the 20 possible κ alleles in these cells has been retained and even this exception is rearranged. In contrast, λ constant region genes remain in the germ-line configuration in each of the eight κ light-chain-producing human B lymphocytes examined. Such observations suggest a hierarchy of light-chain-gene rearrangement beginning with κ and proceeding to λ .

A NUMBER of choices are involved in the differentiation of an antibody-producing lymphocyte. For example, usually only one light-chain gene is functionally expressed in a given cell, the choice having been made between the two light-chain classes (κ or λ) and their respective alleles¹. In addition, after commitment to the expression of a particular light-chain, the phenotypically excluded class and allele apparently remain silent throughout the functional life of the lymphocyte. These phenomena are called isotypic (κ versus λ) and allelic (one κ allele versus the other) exclusion. Whether they are tightly programmed or stochastic remains an extremely interesting but unanswered question.

Fortunately much of what we know about the formation of an active mouse κ gene is relevant to the question of allelic exclusion. In some B (antibody-producing) cells, the silent allele remains in the germ-line configuration and has not undergone the somatic V/J rearrangement necessary to form an active gene^{2–4}. In other cells, the silent allele has undergone V/J recombination but, because of the variability of this reaction, produces an out-of-phase J sequence and, consequently, a nonsense gene^{5,6}. In still other cells, recombination occurs at other than a J region making it impossible to form a coherent coding sequence^{7,8}. In each of these instances, the molecular basis for the silence of the excluded allele is clear, but the basis for the selection and control of this process, if any, remains obscure.

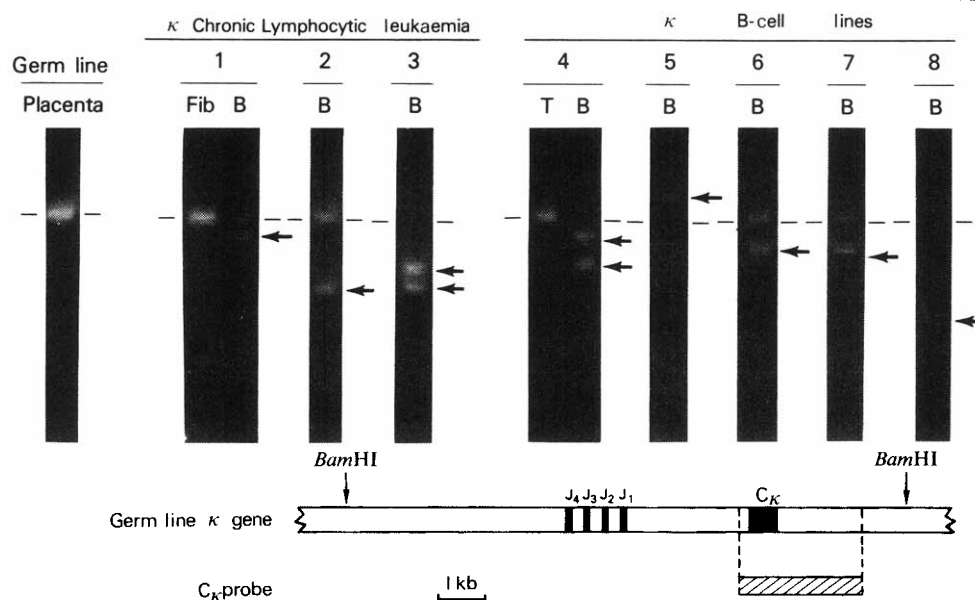
Little is known of the nature of isotypic exclusion, because studies of λ/κ expression in the mouse are necessarily limited by the fact that the λ class is expressed at low frequency in this organism. Recently, however, Alt *et al.*⁹ examined three long-

established λ chain-producing mouse myeloma lines and found that κ genes had been rearranged in two of them and possibly lost in a third. This observation led them to a proposal (see below) regarding the control of allelic and isotypic exclusion by the formation of functional genes. This proposal can be usefully pursued in man. In contrast to the mouse, human immunoglobulins are composed of one-third λ light-chain-bearing molecules and, therefore, the human λ gene system is likely to contribute significantly to antibody diversity. Accordingly, we have turned to man and an array of established B-cell lines and freshly obtained leukaemic cells to ask whether the structure and arrangement of these genes offer some clue as to the mechanism or order of their selection.

We show here that at least one light-chain gene in each light-chain-producing B cell has undergone rearrangement while its allelic partner has either remained in the germ-line configuration, undergone a second rearrangement or been lost. In each cell expressing a κ gene, the λ genes are present in their germ-line configuration, but in surprising contrast, in each λ -producing cell, κ genes have been either deleted or undergone a rearrangement. This consistent correlation suggests, among other possibilities, a putative hierarchy of light-chain-gene rearrangements in which the process is ordered so that it usually progresses from the κ to the λ genes.

κ genes are rearranged and deleted in κ -expressing cells

The arrangement of the κ light-chain genes in human κ -expressing B cells was determined using a radioactive DNA



autoradiograms were developed. The germ-line (embryonic) κ gene map demonstrating the 12-kilobase *Bam*HI C_{κ} -containing fragment and the 2.5-kilobase *Eco*RI embryonic C_{κ} -containing fragment (hatched bar) which was used as the C_{κ} probe are shown. The four human J_{κ} gene segments identified by heteroduplex analysis¹⁰ are indicated.

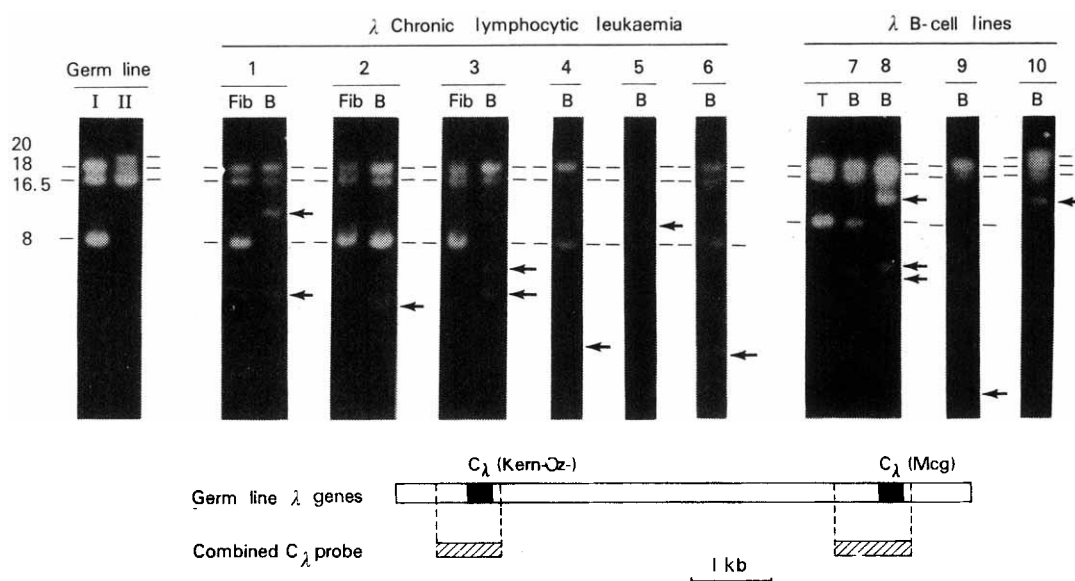
fragment containing the human κ constant region gene¹⁰ as a hybridization probe to detect this gene among fragments generated by *Bam*HI endonuclease digestion of genomic DNA (Fig. 1). Non-rearranged genomic DNA yields a single *Bam*HI fragment containing the single C_{κ} and four human J_{κ} segments¹⁰. Examination of such *Bam*HI fragments from non-B cells (placental tissue, fibroblasts or T cells) derived from 24 different individuals (see representative controls in Figs 1 and 4) showed them all to be in the germ-line configuration. Human κ light-chain-expressing B cells (including lymphocytes derived from three high-count chronic lymphocytic leukaemia patients, three B-cell lines established spontaneously from κ -expressing B-cell malignancies, and two κ -secreting Epstein-Barr virus-transformed human B-cell lines) were then examined. All eight κ -expressing B cells demonstrated at least one *Bam*HI fragment rearrangement when compared with the germ-line configura-

tion (Fig. 1). The second C_{κ} allele remained in the embryonic configuration in four cases, showed a DNA rearrangement in two cases and was deleted in two cases.

λ genes are rearranged in λ -expressing B cells

Analyses of human myeloma proteins¹¹ demonstrated that the λ constant region consists of at least four non-allelic forms in man which bear the non-allelic markers, Kern⁺Oz⁺, Kern⁺Oz⁻, Kern⁻Oz⁻ and Mcg. To identify these genes in λ chain-producing cells, we used a C_{λ} probe consisting of a mixture of two ³²P-labelled DNA fragments derived from cloned human λ constant genes (P.A.H., unpublished results). These were a 1.2-kilobase *Bam*HI-*Eco*RI fragment bearing the C_{λ} (Kern⁺Oz⁻) gene and a 0.8-kilobase *Bam*HI-*Hind*III fragment bearing the C_{λ} (Mcg) gene (Fig. 2, bottom). Because of the close amino acid sequence

Fig. 2 Detection of rearranged λ genes in λ -expressing human B cells. λ -Expressing B cells examined included the monoclonal λ light-chain-bearing B lymphocytes from patients (cases 1-6) with high-count chronic lymphocytic leukaemia, as well as four long-term λ -secreting Epstein-Barr virus-transformed normal human B-cell lines (cases 7, 8, 9, 10). Fibroblast cell cultures from cases 1, 2, 3 and spontaneous T-cell lines from cases 7 and 8 served as matched controls in the germ-line configuration. The two observed germ-line patterns of C_{λ} gene fragments are presented and referred to as type I and type II. As before, genomic DNA from all cells was digested with *Eco*RI, size fractionated by agarose gel electrophoresis, transferred to DBM paper and hybridized with ³²P-labelled combined C_{λ} probe. The C_{λ} probe (hatched bars) was a combination of the 1.2-kilobase embryonic *Bam*HI-*Eco*RI fragment containing the Kern⁺Oz⁻ C_{λ} gene and the 0.8-kilobase embryonic *Bam*HI-*Hind*III fragment containing the Mcg C_{λ} gene (as evidenced by nucleotide sequence analysis; P.A.H., unpublished results).



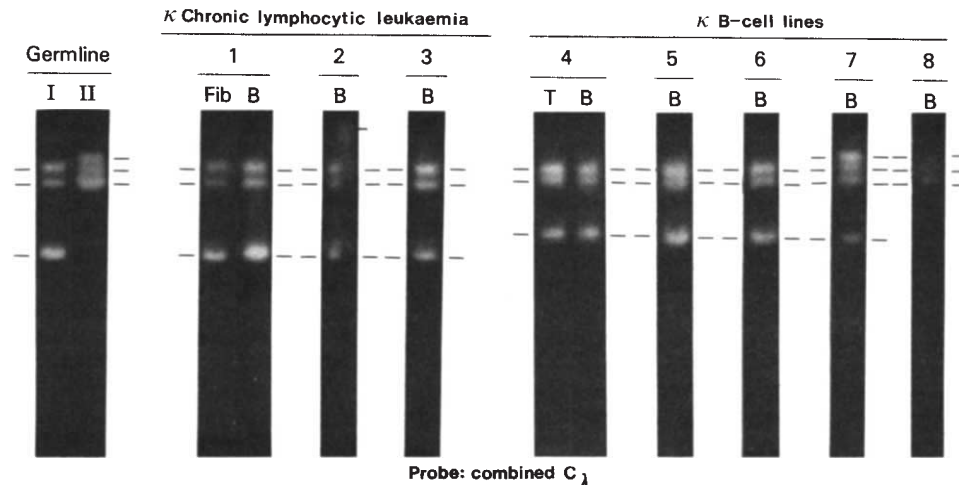


Fig. 3 Detection of λ genes in germ-line configuration in κ -expressing B cells. DNA obtained from the eight κ -expressing B cells shown in Fig. 1 and the available matched germ-line controls were digested with *EcoRI*, size fractionated by agarose gel electrophoresis, transferred to DBM paper and hybridized with the combined C_λ probe (described in Fig. 2). Cases 1–6 show type I germ-line λ genes; case 8, type II germ-line λ genes. We presume that case 7 is heterozygous for both the 8- and 20-kilobase *EcoRI* site alleles, although no matched germ-line control was available for this B-cell line.

homology of the known human λ constant regions, the combined C_λ probe would be expected to cross-hybridize to all four of the C_λ genes. Analyses of *EcoRI* digests of 24 non-B-cell genomic DNAs (placenta, fibroblast and T cells), which were carried out to establish the germ-line pattern, demonstrated a genetic polymorphism in *EcoRI* endonuclease sites that produced two predominant fragment patterns (Fig. 2, germ-line lanes), one which yields 8-, 16.5- and 18-kilobase fragments and another which yields 16.5-, 18- and 20-kilobase fragments. Such polymorphisms, which seem quite common at this locus (P.A.H., unpublished), provide useful chromosomal markers for further genetic studies.

Keeping in mind that apparent alterations in the *EcoRI* fragment pattern at the λ locus might be due to these polymorphisms, we examined a number of human λ chain-expressing B lymphocytes. These included monoclonal B cells from six patients with chronic lymphocytic leukaemia and four normal human B-cell lines transformed by Epstein-Barr virus. *EcoRI* digests of all 10 λ -expressing B cells showed at least one rearranged C_λ allele that was not seen in the matched non-B cells (either fibroblasts or T cells) from the same individual. Three of the B cells displayed double rearrangements that could be used to establish the order of the λ genes in chromosomal DNA. This was done by noting that in three cases (1, 3 and 8; Fig. 2), the 8-kilobase *EcoRI* fragment was lost in a double rearrangement of λ alleles, while the 16.5- and 18-kilobase fragments were preserved. Because V/J recombination occurs to the 5' side of an expressed constant region and involves deletion of DNA between V and J regions^{12,13,19}, the retained constant λ genes must be 3' to those on the 8-kilobase fragment.

λ genes are in the germ-line configuration in κ -expressing B cells

The arrangement of the λ light-chain genes in the eight human κ light-chain-expressing B cells analysed above is shown in Fig. 3. In all eight of the κ -expressing B cells, the *EcoRI* λ gene-

containing fragments remain in the germ-line configuration. Note that case 7, which demonstrates both the 8- and 20-kilobase *EcoRI* restriction fragments, seems to be heterozygous for the type I and type II polymorphic gene configurations.

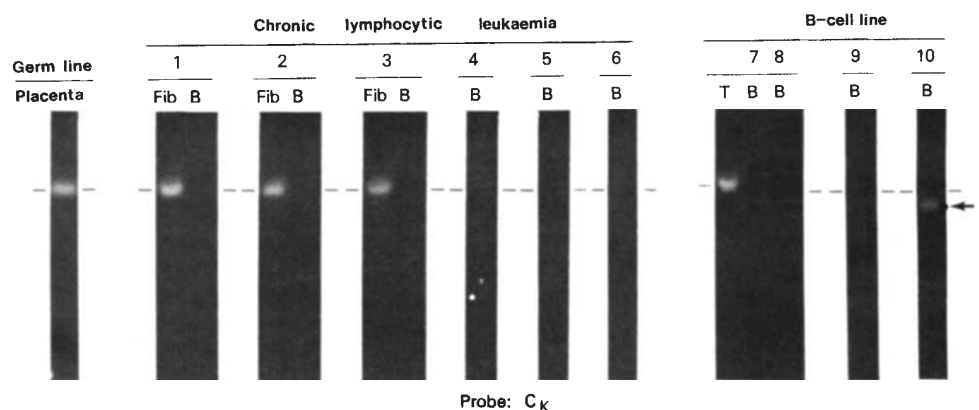
κ genes are deleted or rearranged in λ -expressing B cells

In clear contrast to the germ-line arrangement of λ genes in κ -expressing cells, both C_κ alleles are deleted in all six λ -expressing leukaemic B cells (*BamHI* digest shown in Fig. 4). In addition, both C_κ alleles were deleted in three of the four λ -expressing B-cell lines. In the other case, one κ allele was rearranged while the other was deleted (case 10, Fig. 4). Although case 10 was a long-term cell line, it was not physically cloned. The finding of C_κ gene deletions in the nine λ -expressing B cells was confirmed in the *EcoRI* digests of genomic DNA from these same cells (data not shown). In contrast to the observed deletion in the clonal B cells, the C_κ gene was present in the germ-line configuration in *BamHI* and *EcoRI* digests of DNA from matched (same individual) T cells or fibroblast lines, demonstrating that the κ gene deletions are apparently limited to the λ -expressing B cells (Fig. 4).

The κ gene deletions usually include the J segment

To examine the possibility that the C_κ deletion could extend to J and V region coding segments, conceivably involving the entire chromosome, we prepared probes corresponding to four J_κ segments and to a V-region gene (see Fig. 5) and tested for complementary sequences in each cell line. In the representative case shown (Fig. 5), C_κ and J_κ coding segments fail to hybridize to a complementary sequence within the λ -expressing B cell, indicating that the deletion has included the J_κ segment. However, the V-region probe detects complementary fragments indicating that this V region family has not been included in the deletion of C_κ and J_κ . The presence of V_κ sequences demon-

Fig. 4 C_κ gene deletion or rearrangement in λ -expressing B cells. The 10 λ -expressing B cells (B) and their available matched germ-line control fibroblasts (Fib) or T cells (T) were *BamHI*-digested as before and then hybridized with the ³²P-labelled 2.5-kilobase *EcoRI* C_κ gene fragment (see Fig. 1). We see very faint bands present in several cases of chronic lymphocytic leukaemia B cells that are not evident in the figure. These represent the germ-line C_κ segments present in the 3–15% contaminating monocytes and T cells found in leukaemic B-cell preparations and are absent from all the cultured cell lines.



strates that the deletion event is localized and does not represent merely a loss of the chromosome bearing the κ locus. Identical studies of the remaining deletion examples gave identical results with a single exception—in one case, a J_{κ} was present, but in a rearranged configuration (data not shown). This exception nevertheless suggests that such deletions can occur from various points in the chromosomal DNA.

Several gene configurations can account for allelic and isotype exclusion

It is already clear from various analyses in the mouse that a significant proportion of κ light-chain-gene rearrangements are aberrant; that is, they create non-functional or invalid genes, genes incapable of producing normal light-chains⁵⁻⁸. Multiple rearrangements are not uncommon in mouse myeloma cells^{5-8,14-18}. Some of these rearrangements seem to depend on the normal V/J recombination mechanism that occasionally creates a nonsense sequence and, hence, a null gene^{5,6}. Other rearrangements do not seem to depend on V/J recombination, but involve additional mechanisms (conceivably those analogous to the heavy-chain switch) that produce aberrant recombinants between regions that may or may not encode authentic V or J segments^{7,8}. One such rearrangement actually involves a

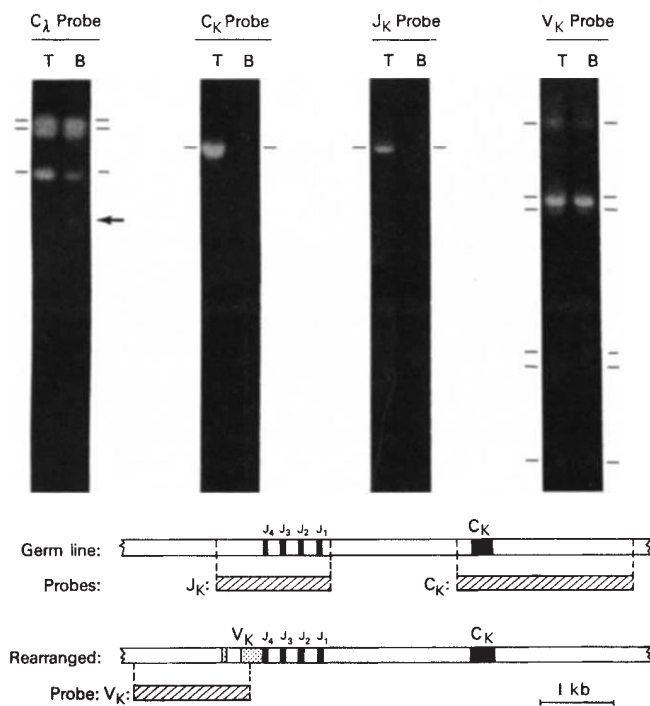


Fig. 5 Extent of κ gene deletion in λ -expressing B cells. A comparison of C_{λ} , C_{κ} , J_{κ} and V_{κ} gene segments in a representative λ -expressing B cell (case 7, Fig. 2) and its matching (same individual) T cell is shown. The probes are shown as hatched bars under each map. ^{32}P - C_{λ} and ^{32}P - V_{κ} 6410 (see below) probes (shown in figure) were hybridized to the *Eco*RI digests of B cells and available germ-line controls. ^{32}P - J_{κ} and ^{32}P - C_{κ} probes were hybridized with the *Bam*HI digests of the same cells. The V_{κ} gene probe was the 1.6-kilobase *Hpa*II fragment derived from the active V_{κ} - J_{κ} recombinant of the κ -expressing cell line RPMI 6410¹⁰. This fragment includes 63 bases of the 5' variable region corresponding to the first framework (sequence not shown). The amino acid translation of the nucleotide sequence yielded a protein sequence identical to that obtained for the human light-chain MAG²³ which is a member of κ subgroup II²⁴. The J_{κ} gene probe was a 1.7-kilobase germ-line (embryonic) *Cfo*I restriction fragment containing four human J_{κ} segments. The C_{κ} probe is described in Fig. 1 legend; the C_{λ} probe, in Fig. 2 legend.

recombinational 'hotspot' between the mouse J and C regions that is complementary to an extension of the small palindromic sequence thought to be involved in V/J recombination⁷. It is not known what proportion of DNA rearrangements during B-cell development result in the formation of an active gene, but it is conceivable that all light-chain alleles of both classes will occasionally rearrange in a single cell to form non-functional genes. Such cells may not survive as expanded clones. However,

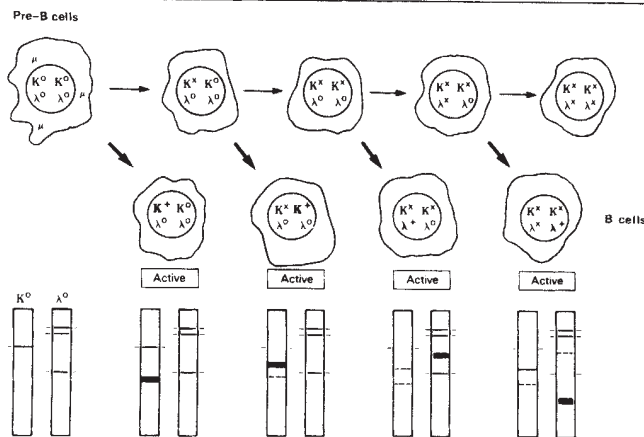


Fig. 6 Developmental scheme of ordered light-chain gene rearrangements in pre-B cells. The genotype of each intermediate cell is indicated within the nucleus, while the predicted arrangement of *Bam*HI(κ) or *Eco*RI(λ) fragments is shown diagrammatically below each active cell. The overall scheme of lymphocyte development is not shown in order to focus on those stages concerned with light-chain gene activation. The initial cell in this scheme represents a pre-B cell already committed to the expression of cytoplasmic μ heavy chain. The expected germ-line configuration of the κ (κ^0) and λ (λ^0) light-chain gene fragments is shown schematically below. Light-chain gene rearrangements then take place (either productively or aberrantly) in a sequence such that rearrangements progress generally from the κ to the λ light-chain genes. Gene rearrangements cease to occur as a result of the formation of an active light-chain gene and synthesis of a viable light-chain product as suggested by Alt *et al.*⁹. The genotypes and representative light-chain gene fragment configurations for each of the four light-chain-producing cell types are shown. Bold faced lines represent the active allele; a dotted line represents either a deleted or aberrantly rearranged allele. Genotype superscripts signify the following: 0, germ-line configuration; +, viable V-J joining event; X, aberrant rearrangement or deletion. Gene fragment patterns consistent with each active cell type have been noted in this study. In addition, the scheme predicts a population of B-cell precursors with non-rearranged or only one aberrantly rearranged gene. It also predicts a population of cells incapable of forming valid light-chains, having 'spent' their germ-line genes on invalid recombinations.

in a B cell in which a functional light-chain gene has been formed, it is reasonable to suggest that germ-line or aberrantly rearranged genes account for allelic exclusion. Perry *et al.*⁴ have emphasized that non-viable gene rearrangements may have a significant role in this process, whereas Joho and Weissman³ have interpreted their studies of κ gene arrangements in circulating lymphocytes so as to suggest that the excluded allele is most often in the germ-line configuration. Similar arguments can be made to account for isotype exclusion. The present study suggests that both aberrantly rearranged genes and genes that remain in the germ-line configuration are responsible for these types of exclusion.

The order of light-chain gene rearrangements

The fact that λ genes seem to remain in the germ-line configuration in κ -producing cells while κ genes are rearranged or (more often) deleted in λ -producing cells allows us to suggest a hierarchy of events involved in light-chain activation. If we simply assume that κ gene rearrangement usually precedes that of λ , and if we accept the useful proposal of Alt *et al.*⁹ that the appearance of a functional light-chain product precludes additional light-chain-gene rearrangements, we can place each of our B cells along the developmental scheme shown in Fig. 6. In this scheme it is proposed that a pre-B cell undergoes either a valid or a non-functional κ rearrangement. If the initial rearrangement is valid, the clone expands and the gene arrangement $\kappa^+/\kappa^0/\lambda^0/\lambda^0$ (where the superscripts + = valid, 0 = germ-line, X = aberrantly rearranged or deleted) is observed. If the initial κ rearrangement is non-functional, a second κ rearrangement occurs and, if valid, the clone expands and the arrangement $\kappa^X/\kappa^+/\lambda^0/\lambda^0$ is observed. If both κ alleles are aberrantly rearranged or deleted, the next rearrangement involves one of the λ alleles which, following the cascade diagram shown in Fig. 6, could go on to produce all the possible combinations of valid and non-functional genes shown in the figure. An intermediate population of pre-B cells which have no active light-chain genes

is predicted by this scheme, including a terminal cell type in which all light-chain genes are aberrantly rearranged or deleted. Each active configuration predicted by this scheme occurs among the B cells we have examined. On the other hand, if λ rearrangement were to precede that of κ or to occur at the same relative frequency as κ , then we should have seen λ genes rearranged in κ -producing cells. We have not observed this situation in any of the κ -producing cells examined. A sequential process of light-chain gene rearrangement, which seems to involve the frequent deletion of C_κ genes and a joining order that proceeds from κ to λ , can be thought of as an error-compensating cascade that finally creates an active gene while experimenting with the flexible joining rules that allow optimum diversification.

Why is the κ constant region deleted?

The arguments outlined above do not deal with the finding that the κ constant region genes are entirely deleted in 9 of the 10 λ chain-producing cells analysed. Such deletions obviously involve recombination sites not ordinarily involved in V/J recombination. For example, the deletion of C_κ may be the result of an aberrant recombination event between a V_κ gene and a recombination site which lies 3' to the κ constant gene. Alternatively, this deletion might represent secondary recombinational events that follow the formation of valid or invalid κ genes. But why has the gene disappeared with such a high frequency? One possibility is that this deletion is a normal step in

the progress of light-chain recombination, perhaps a step for the elimination of inactive κ genes. This seems unlikely in view of the fact that cells with active κ genes seem to tolerate the presence of invalid allelic recombinants. It is more likely that elimination of the κ gene is closely related to λ gene rearrangement, either as an effect or as a cause. The possibility that λ rearrangement is necessary for κ gene deletion seems doubtful in view of the several examples of κ deletion in κ chain-producing B cells in which no λ gene rearrangement has occurred (Fig. 1, cases 5 and 8). Instead, one might imagine the formation of a positive control element by the κ deletion event which causes the onset of λ gene rearrangements. Alternatively, a negative control element preventing λ gene rearrangement might be eliminated by the κ deletion event. Our single example (Fig. 4, case 10) of a retained (but rearranged) κ gene in a λ -producing B-cell line does not rule out the negative control argument as the effect might be sensitive to gene dosage or the κ gene rearrangement may itself have been sufficient to eliminate the control element. In addition, because this human cell line is uncloned, we must consider the possibility that the rearranged κ gene might reside in a subpopulation of B cells that is not involved in λ chain production. The striking observation that κ constant region genes are consistently deleted or rearranged in λ -expressing B cells may, among other possibilities, indicate a potential coupling between λ chain activation and κ gene deletion, a recombination event that might have regulatory implications.

Received 13 November 1980; accepted 14 January 1981.

1. Melchers, F., von Boehmer, H. & Philips, R. A. *Transplant Rev.* **25**, 26–58 (1975).
2. Seidman, J. G. & Leder, P. *Nature* **276**, 790–795 (1978).
3. Joho, R. & Weissman, I. *Nature* **284**, 179–181 (1980).
4. Perry, R. P. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 1937–1941 (1980).
5. Max, E. E., Seidman, J. G., Miller, H. & Leder, P. *Cell* **21**, 793–799 (1980).
6. Altenburger, W., Steinmetz, M. & Zachau, H. G. *Nature* **287**, 603–607 (1980).
7. Seidman, J. G. & Leder, P. *Nature* **286**, 779–783 (1980).
8. Leder, P. *et al. Cold Spring Harb. Symp. quant. Biol.* (in the press).
9. Alt, F. W., Enea, V., Bothwell, A. L. & Baltimore, D. *Cell* **21**, 1–12 (1980).
10. Hlieter, P. A., Max, E. E., Seidman, J. G., Maizel, J. V. & Leder, P. *Cell* **22**, 197–207 (1980).
11. Felt, J. & Deutsch, H. *Immunochimistry* **12**, 643 (1975).
12. Sakano, H., Huppi, K., Heinrich, G. & Tonegawa, S. *Nature* **280**, 288–294 (1979).

13. Seidman, J. G. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 6022–6026 (1980).
14. Lenhard-Schuller, R., Hohn, B., Brack, C., Hiram, M. & Tonegawa, S. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4709–4713 (1978).
15. Wilson, R., Miller, J. & Storb, U. *Biochemistry* **18**, 5013–5021 (1979).
16. Steinmetz, M., Zachau, H. G. & Mach, B. *Nucleic Acids Res.* **6**, 3213–3229 (1979).
17. Steinmetz, M. & Zachau, H. G. *Nucleic Acids Res.* **8**, 1693–1707 (1980).
18. Wilson, R., Miller, J. & Storb, U. *Biochemistry* **18**, 5013–5021 (1979).
19. Max, E. E., Seidman, J. G. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3454–3460 (1979).
20. Polsky, F., Edgell, M. H., Seidman, J. G. & Leder, P. *Analyt. Biochem.* **87**, 397–410 (1978).
21. Alwine, J. C. *et al. Meth. Enzym.* **68**, 220–242 (1980).
22. Rigby, P. J., Dieckmann, M., Rhodes, C. & Berg, P. *J. molec. Biol.* **113**, 237–251 (1977).
23. Sletten, K., Hannestad, K. & Harbol, M. *Scand. J. Immun.* **3**, 215–218 (1974).
24. Kabat, E. A., Wu, T. T. & Bilofsky, H. *NIH Publ. No.* 80–2008, p. 10 (1979).

Aberrant rearrangements contribute significantly to the allelic exclusion of immunoglobulin gene expression

Christopher Coleclough, Robert P. Perry, Klaus Karjalainen & Martin Weigert

Institute for Cancer Research, Fox Chase Cancer Center, Philadelphia, Pennsylvania 19111, USA

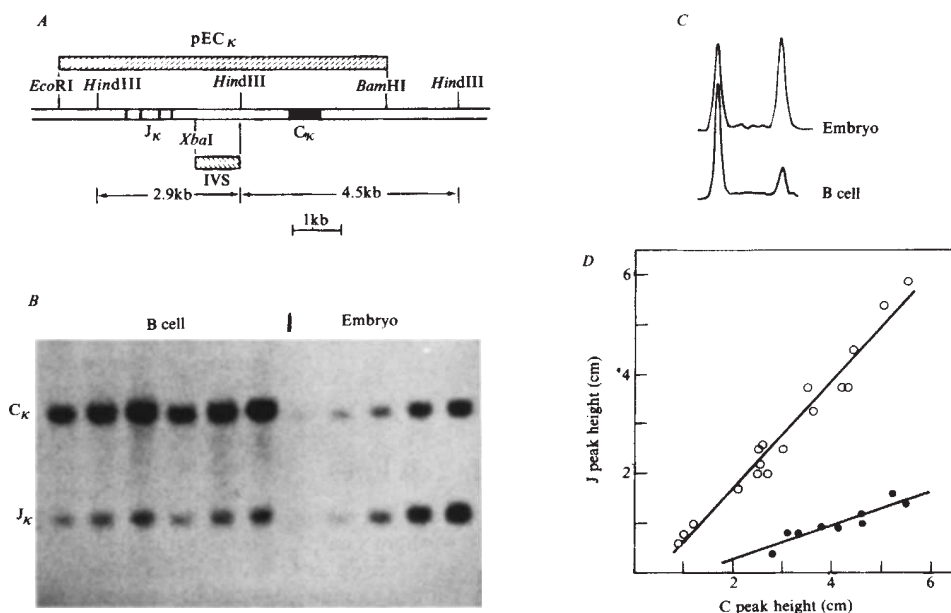
A study of the organization of light- and heavy-chain immunoglobulin genes in mouse splenic B cells, spleen-derived hybridomas and plasmacytomas has unequivocally demonstrated that aberrant rearrangements are common during normal B-cell development. The results support a probabilistic model for allelic exclusion of immunoglobulin gene expression.

PHENOTYPIC expression of immunoglobulin genes depends on site-specific recombination events which bring the genetic elements encoding the immunoglobulin variable and constant regions within a single transcription unit. This is most easily appreciated in the mouse λ_1 light-chain gene system, in which an active gene is constructed by the precise fusion of the single V_{λ_1} gene with the single J_{λ_1} element residing 1.2 kilobases 5' of the C_{λ_1} gene¹. Although the mouse κ light-chain gene family is more complex, comprising some hundreds of V_κ genes and four functional J_κ elements, gene activation is essentially similar, requiring a single recombination event to fuse V_κ and J_κ ^{2,3}.

However, the construction of a functional μ heavy-chain gene seems to require two recombination events, which result in the fusion of three distinct genetic elements: V_H , D_H and J_H ^{4,5}.

Immunoglobulin gene expression exhibits allelic exclusion; that is, in each individual antibody-forming cell, only one of the pair of alleles at an immunoglobulin locus is expressed while the other is phenotypically silent. Similarly, the synthesis of κ and λ light chains in individual lymphoid cells is mutually exclusive. How is this achieved? Recently, we analysed the status of allelic κ genes in a large number of κ -secreting plasmacytomas⁶. In about half of the plasmacytomas examined, the phenotypically

Fig. 1 The extent of J_{κ} rearrangement in κ -bearing B cells. **A**, schematic map of the J_{κ} - C_{κ} region. Cloned hybridization probes are shown as hatched boxes. Plasmid pEC $_{\kappa}$ was constructed by insertion into pBR322 of the *Eco*RI-*Bam*HI fragment shown, which was derived from the 15.5-kilobase (kb) *Eco*RI fragment cloned in Charton 4A by Lenhard-Schuller *et al.*²³. IVS was isolated by *Hind*III-*Xba*I digestion of pEC $_{\kappa}$ and gel electrophoresis. Recombination involving the J_{κ} region will deplete the germ-line J_{κ} -containing *Hind*III fragment relative to the C_{κ} -containing fragment, and may be detected as a decrease in the J_{κ} / C_{κ} fragment ratio. Recombination 3' (right) of the *Hind*III site between J_{κ} and C_{κ} , however, is not detectable by this assay, as it depletes the germ-line J_{κ} and C_{κ} *Hind*III fragments to the same extent. **B**, a representative experiment. DNA samples prepared from 14-day BALB/c embryos²⁴, or from κ -bearing B cells were digested to completion with *Hind*III (Biolab), and various amounts (~0.5–8 μ g) electrophoresed through 0.8% neutral agarose gels, transferred to nitrocellulose paper²⁵ (Schleicher and Schuell), and hybridized to ³²P-labelled pEC $_{\kappa}$ plasmid DNA^{11,26}. Filters were exposed to X-ray film at -70 °C, using an intensifying screen, for various periods; the one shown was exposed for ~10 h. The κ -bearing B cells were purified from BALB/c spleen cells. Suspensions of spleen cells were depleted of T cells by treatment with anti-Thy-1 and complement, and centrifuged through Ficoll-Hypaque (Pharmacia) to remove erythrocytes and dead cells. The B-cell-enriched preparation was labelled with rabbit antiserum directed against mouse κ constant region, counterstained with fluorescent-coupled goat anti-rabbit immunoglobulin and passed through the fluorescence-activated cell sorter (Becton-Dickinson, FACS II). **C**, densitometer tracings of representative tracks from Southern blot autoradiographs of embryo and B-cell DNA showing the marked reduction of the J_{κ} peak (right) in B-cell DNA. **D**, graphical estimation of the extent of J_{κ} rearrangement in B cells. Several exposures of the autoradiogram in **B** and another set of data from a similar experiment were scanned as in **C**. Comparison of multiple exposures of the same tracks showed that peak heights between 0.5 and 6 cm corresponded to the linear portion of the dose-response curve of the film. Data within this range were plotted as J_{κ} against C_{κ} ; each point is derived from an individual track. The least-squares best fit of the embryo data has a slope of 1.105; that of the B-cell data, 0.35. ○, Embryo DNA; ●, B-cell DNA.



silent allele was maintained in the germ-line context, whereas in the other half it had undergone some type of non-productive rearrangement. We concluded that imprecision intrinsic in the process of somatic recombination contributes significantly to the allelic exclusion of κ gene expression, and we proposed a model in which the status of the κ alleles is determined by the recombination frequency and the relative probabilities of productive as opposed to non-productive rearrangements^{6,7}. Our conclusion, which was supported by the results of other studies on the extent of J_{κ} ^{23,29,30} and J_{H} ³¹ rearrangement in myelomas and lymphomas, was based on the assumption that plasmacytomas are valid models of untransformed plasma cells. This assumption was questioned by Joho and Weissman⁸, who measured the extent of κ gene rearrangement in κ -expressing splenic B cells and concluded that non-productively rearranged κ genes are very rare or absent in normal lymphoid populations. However, their conclusion was based on a single determination and an arbitrary interpretation of their normalization data.

To settle this issue and to examine the general applicability of the probabilistic model we have made a detailed study of the extent of non-productive rearrangement of both light- and heavy-chain immunoglobulin genes in normal mouse splenic B cells, spleen cell-derived hybridomas and, for comparison, appropriate panels of plasmacytomas and B-cell lymphomas. Our results indicate that at least one out of three normal B cells contains a non-productively rearranged J_{κ} locus and that the silent J_{H} locus is rearranged in >80% of the B-cell population. We also provide evidence for extensive J_{κ} rearrangement in Λ -producing cells in contrast to a paucity of J_{Λ} rearrangement in κ expressors. These results demonstrate unequivocally that aberrant rearrangements are a common feature of normal B-cell development. Moreover, they support the idea that both class and allelic exclusion can be explained in terms of a probabilistic model of immunoglobulin gene rearrangement.

Non-productive rearrangement of κ genes occurs frequently in B cells

In monoclonal lymphoid lines, J_{κ} rearrangement is detected by observing a discrete J_{κ} -containing restriction fragment differing

from that found in embryo (germ-line) DNA. In contrast, one would not expect a discrete novel J_{κ} band in a Southern blot of DNA from the total splenic population of κ -expressing B cells because the many different V-J combinations would result in a diverse array of fragments of varying sizes. The extent of J_{κ} rearrangement in this population can, therefore, be calculated from the relative decrease in intensity of the embryo-type J_{κ} band. We needed a strategy to estimate precisely this decrease, using controls and quantitative measurements internal to the experiment. Figure 1A shows a map of the C_{κ} region. It can be seen that a *Hind*III digest of embryo DNA when probed on a Southern blot with a plasmid containing an appropriate *Eco*RI-*Bam*HI fragment, pEC $_{\kappa}$, reveals two major fragments. The larger of these, ~4.5 kilobases, contains the C_{κ} gene, while all the J_{κ} elements are contained within a fragment of ~2.9 kilobases. As these fragments overlap the pEC $_{\kappa}$ probe to about the same extent, their autoradiographic intensities should be roughly equivalent. That this is indeed the case was verified (Fig. 1B-D) by measurements of varying amounts of embryo DNA. As shown in Fig. 1D, a plot of the intensity of the J_{κ} band versus that of the C_{κ} band gives a straight line with a slope of ~1. For these determinations, autoradiographic exposure times were confined to a range in which the densitometric peak heights were directly proportional to the amount of radioactivity emitted. In contrast, when a similar set of measurements were made with the DNA of κ -expressing splenic B cells, the intensity of the J_{κ}

Table 1 Extent of J_{H} gene rearrangement in splenic B cells

Sample	C_{μ} gene dosage (μ g embryo DNA equivalent)	J_{H} gene dosage (μ g embryo DNA equivalent)	% J_{H} genes unrearranged	% B cells with both J_{H} genes rearranged*
B1	10.0	~0	~0	~100
B2a	8.0	0.5	6.25	87
B2b	7.5	0.7	9.3	81

The C_{μ} and J_{H} gene dosages were estimated from Fig. 5B.

* Assumes each cell has at least one rearranged J_{H} .

band was ~ 0.35 that of the C_κ band as judged by the slope of the J_κ versus C_κ plot. If all the B cells had contained one productively arranged and one unrearranged J_κ locus, the slope would have been 0.5. Our data demonstrate that this is not the case (confidence limit >0.95); rather they indicate that about one-third of the cells contain two rearranged J_κ loci, one of which would presumably be non-productive with respect to κ -chain production. This fraction is actually a minimum estimate because the above analysis is sensitive only to those rearrangements which result in a loss of the 2.9-kilobase J_κ fragment and retention of the 4.5-kilobase C_κ fragment, whereas some non-productive rearrangements could conceivably involve the loss of both fragments.

λ_1 genes are infrequently rearranged in κ -expressing cell lines but κ genes are often rearranged in λ -producing cell lines

We wished to assess whether the probabilistic model which we applied to allelic exclusion at κ loci is also valid for light-chain class exclusion. That is, can both κ and λ loci be regarded as elements competing for mutually exclusive expression in a single pool of B cells? If this is the case, then non-productive gene rearrangement would be evident in cell lines producing either κ or λ light chains.

Initially we determined the status of λ_1 genes in a panel of κ -producing plasmacytomas. Southern blots of *Eco*RI-digested embryo DNA probed with a λ_1 cDNA-containing plasmid exhibit four bands at 8.6, 4.8, 3.5 and 2.5 kilobases (Fig. 2). Brack *et al.*¹ have identified the largest three as containing C_{λ_1} , $V_{\lambda_{H1}}$ and V_{λ_1} , respectively. We do not know which hybridizing element is contained within the 2.5-kilobase band; possibly it is related to the λ_{H1} locus⁹. Each of the four λ_1 -producing plasmacytomas examined showed a novel band at 7.4 kilobases, characteristic of the fused V_{λ_1} - C_{λ_1} gene¹. There is some evidence of additional λ gene rearrangement: Y5444 seems to lack an unrearranged allelic C_{λ_1} gene and J558 lacks an unrearranged

allelic V_{λ_1} gene. When DNA from 14 κ -producing plasmacytomas was similarly probed, all revealed the germ-line set of λ bands; one cell line, PC2960, showed one additional band that is different from the productive V_{λ_1} - C_{λ_1} hybrid. One λ gene has been rearranged non-productively in this cell line, but without further analysis, we cannot say which. Clearly, however, λ gene rearrangement is rare in κ -producing cells.

We performed the complementary experiment, determining the status of κ genes in λ -producing cells, at first with the four λ plasmacytomas used above. In no case could we find a germ-line-type κ gene (not shown). To extend our analysis, we decided to use λ -producing hybridomas, which can be produced in reasonable numbers and, moreover, can be assayed soon after induction, minimizing the time spent in serial passage. The use of cultured hybridoma lines also eliminates any possible contamination by DNA of non-lymphoid origin. Certain strains of mice, in response to primary immunization with nitrophenyl chicken globulin (NPCG), produce antibodies in which λ light chains are frequent¹⁰. We produced a large number of hybridomas from spleen cells after primary immunization of CB.17 mice with NPCG. After cloning and serological characterization, we analysed DNA from 12 λ -secreting hybridomas. DNA was digested with *Hind*III and processed as above, this time probing with the *Hind*III-*Xba*I fragment, IVS. As shown in Fig. 1A, embryo DNA yields a 2.9-kilobase fragment that hybridizes with this probe; rearrangement at or close to any J_κ will substitute this with a novel fragment. Our results are shown in Fig. 3. Ag8.653, the myeloma partner used to produce hybridomas, has a strong band at ~ 6.5 kilobases, which is maintained in all the hybridomas. Clearly, rearrangement at J_κ is frequent in these hybridomas. Four hybridomas, d, e, f and g, showed bands similar in mobility to that characteristic of unrearranged C_κ genes. A second analysis, including an embryo marker, shows that only the fragment from hybridoma f is actually coincident with the embryo-type fragment (Fig. 3B). To verify that hybridoma f does indeed contain a non-rearranged κ gene, we performed a similar analysis with DNA that was digested with *Bam*RI and *Bam*HI plus *Eco*RI. Hybridoma f continues to

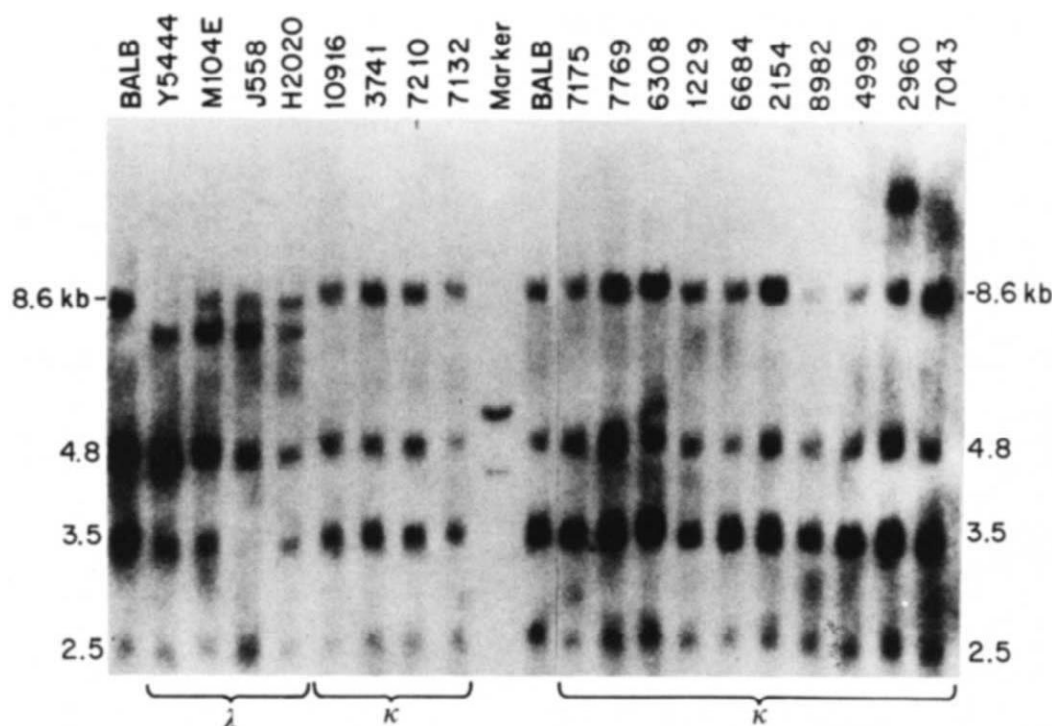
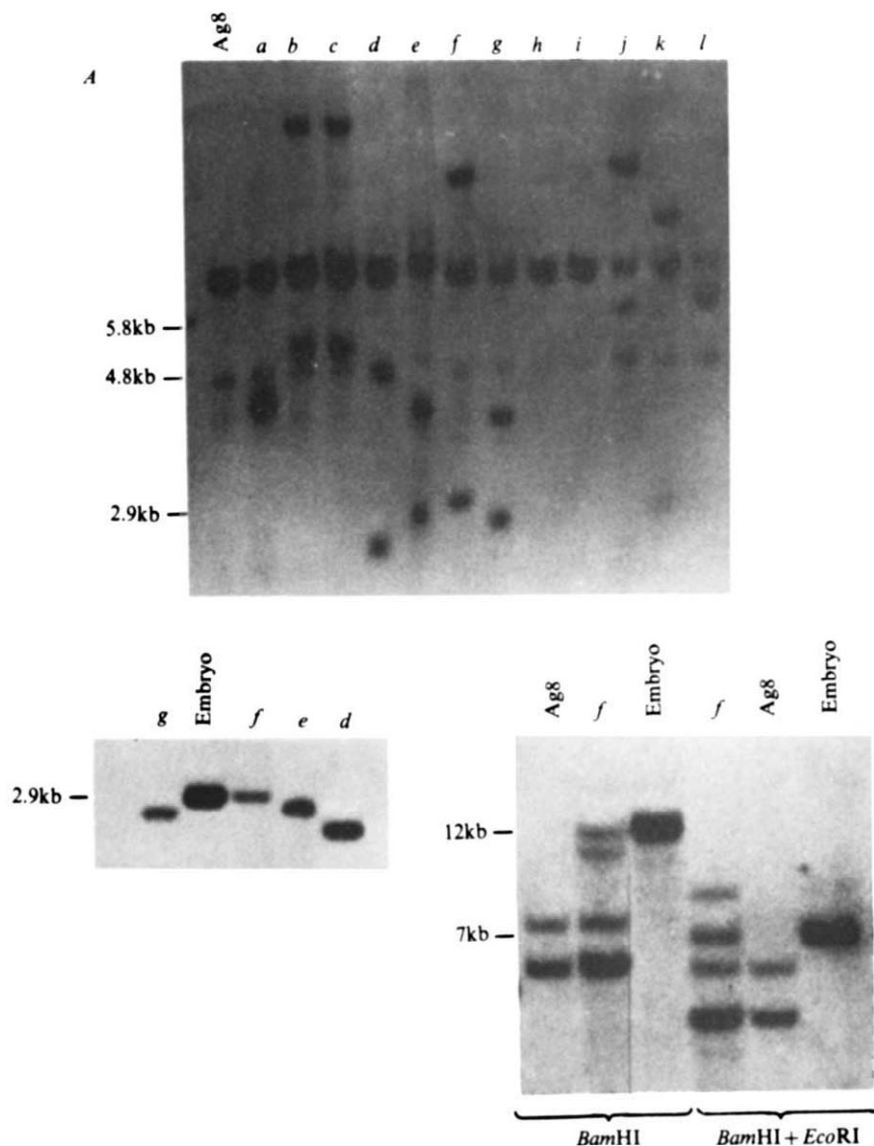


Fig. 2 Southern blot analysis of the λ loci in a series of λ - and κ -secreting plasmacytomas. Aliquots (10 μ g) of DNA from the sources indicated were digested with *Eco*RI, electrophoresed through a 0.7% agarose gel, transferred to nitrocellulose and hybridized with 32 P-labelled plasmid B1, which contains V_{λ_1} and C_{λ_1} sequences derived from H2020 mRNA¹. Tracks marked BALB contained 10 μ g 14-day BALB/c embryo DNA. kb, Kilobase.

Fig. 3 J_{κ} rearrangement in λ -producing hybridomas. Hybridomas specific for NPCG were produced as described elsewhere¹⁰ by fusion of immune spleen cells with Ag8.653 cells²⁷. After selection for drug-resistance phenotype and antigen binding specificity, those hybridomas secreting λ light chains were revealed by radioimmunoassay. **A**, DNA from the myeloma parent cell line Ag8.653 and from each of 12 individual cloned λ -producing hybridomas (a–l) was digested with *Hind*III, electrophoresed through 0.7% agarose and probed with ³²P-labelled IVS (see Fig. 1A). Ag8.653 DNA shows strong hybridization to a fragment ~6.5-kilobases long, and this is maintained in all hybridomas. The weak band at 4.5 kilobases is due to slight contamination of the electrophoretically purified IVS probe with a C_{κ} fragment, which hybridizes to the 4.5-kilobase C_{κ} *Hind*III fragment and thus is maintained in all cell lines. The J_{κ} region in its germ-line context is revealed as a 2.9-kilobase fragment. **B**, four hybridomas, d–g, which exhibited J_{κ} *Hind*III fragments with mobilities close to that characteristic of the germ-line J_{κ} region, were re-assessed relative to BALB/c embryo DNA. *Hind*III digests were analysed by gel electrophoresis and Southern blotting as above. **C**, hybridoma f was further analysed to confirm that it contained a non-rearranged J_{κ} region; f has a fragment with identical mobility to that of the germ-line fragment after both *Bam*HI digestion and *Bam*HI-*Eco*RI double digestion.



display a fragment indistinguishable in size from the embryo-type κ gene fragment (Fig. 3C).

Thus, the prevalence of rearranged over non-rearranged κ genes in λ -producing lymphoid lines is high but not absolute. Of the 15 distinct J_{κ} gene contexts (counting b and c as one clone; see below) discovered in these 12 hybridomas, one is germ line and 14 are rearranged. Three of the hybridomas show only one J_{κ} fragment; two show no hybridization other than to the Ag8 fragment. This could be due either to chromosome segregation or to the specific deletion of this region. Interestingly, two clones, b and c, show identical patterns of J_{κ} context, and are also identical in heavy-chain gene context (not shown). We are certain that these are independent clones; thus they probably derive from the same B cell, clonally expanded *in vivo*.

J_H rearrangement usually occurs in both heavy-chain alleles in plasmacytomas and normal B cells

We set out to evaluate the incidence of rearrangement in the silent J_H locus by examining the context of the J_H cluster in established lymphoid lines and splenic B cells. The J_H elements are contained in an *Eco*RI fragment ~6.2 kilobases long (Fig. 4A). Recombination within or near the J_H cluster should produce a novel *Eco*RI fragment which could be revealed by hybridization of an appropriate recombinant plasmid, pJ11. An application of this analysis to a series of plasmacytomas is shown

in Fig. 4B. As all the plasmacytomas synthesize a heavy chain, at least one novel *Eco*RI fragment is expected in each case. In fact, most, 15 out of 20, show two distinct novel *Eco*RI fragments, as do the two lymphomas 38C and 70Z, demonstrating frequent rearrangement of the silent allelic J_H cluster. All plasmacytomas which were grown as subcutaneous tumours show some, mostly faint, hybridization to an embryo-type fragment which presumably derives from contaminating host cells, as noted before^{6,11}. Rearrangement of both J_H alleles was also observed by Cory and co-workers³¹ in a similar analysis of a myeloma and several lymphoma lines. Three examples, PC2960, PC7043 and PC6308, show strong hybridization to an embryo-type fragment and only a single novel fragment. We have further analysed PC7043 and PC6308 using an *Eco*RI-*Sac*I digest, which yields a J_H fragment of ~3 kilobases. Both plasmacytomas yield a J_H band coincident with that of the embryo type (data not shown). Thus, we note that, although rearrangement of the silent allelic J_H cluster is frequent in lymphoid lines, it is not a universal occurrence.

In our first set of experiments with splenic B cells purified on a fluorescence activated cell sorter (FACS), we noted that, whereas *Eco*RI digestion of B-cell DNA yields a distinct C_{μ} fragment, no discrete band is obtained using *Kpn*I¹². Because the *Kpn* site upstream of C_{μ} is 5' distal to the J_H cluster, this result suggested relatively frequent J_H recombination in B cells. Here we refine and quantify those observations.

As discussed above for J_{κ} rearrangement, the extent of J_H rearrangement in the splenic population of surface IgM-positive

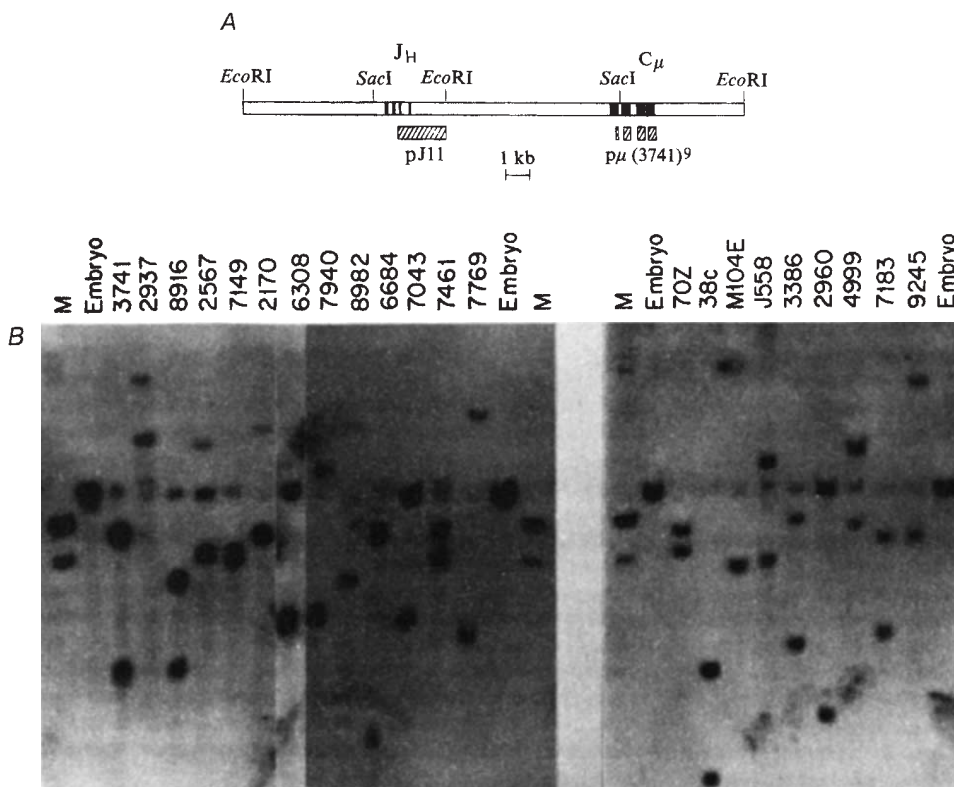


Fig. 4 J_H gene rearrangement in lymphoid lines. **A**, restriction map of C_μ gene and J_H region. Black areas indicate protein coding elements. Hatched boxes indicate sequences represented in the C_μ probe: the cDNA plasmid p μ (3741)⁹ and the J_H probe: the genomic DNA plasmid pJ11¹³. The map was compiled from refs 13 and 28 and our own observations. Not all $SacI$ sites are shown. **B**, 10 μ g of DNA prepared from each indicated source was digested with $EcoRI$ and electrophoresed through 0.75% agarose, transferred to nitrocellulose and hybridized with ^{32}P -labelled pJ11. All cell lines are from NZB mice, except M104E and J558, which are from BALB/c mice, 38C from a C3H mouse, and 70Z which derives from a (C75BL/6 $\times$ DBA/2) F₁ hybrid. The mobility of non-rearranged J_H restriction fragments from these strains is identical, thus all tracks are comparable with the BALB/c embryo tracks. Tracks labelled M contained 1 μ g $EcoRI$ -digested λ phage DNA and 3 pg each of $EcoRI$ -treated, and $EcoRI$ plus $SacI$ -digested pMB9 DNA. The latter species are revealed by hybridization with recombinant plasmid probes and appear on the autoradiograph as 5.8- and 4.8-kilobase fragments.

B cells must be calculated from the relative decrease in intensity of the embryo-type J_H band. We again used hybridization to the constant-region gene fragment to control for and quantify the extent of hybridization to the neighbouring J-region fragment. We wished to separate the C_μ gene from its labile 5' flanking region¹³ to avoid the possibility that any rearrangement in this region could lead to an underestimation of C_μ gene dosage. This was achieved using $SacI$, which cleaves in the intron between $C_{\mu 1}$ and $C_{\mu 2}$ exons. A $SacI$ - $EcoRI$ double digestion yields a 5.1-kilobase C_μ fragment and a 3-kilobase J_H fragment from embryo DNA (Fig. 4A).

Our analysis of splenic B cells is shown in Fig. 5A. Two different samples of B-cell DNA were analysed, one in duplicate, loading ~6 to 8 μ g of DNA per track, together with two sets of a titrated series of embryo DNA. B-cell DNA shows a striking deficit of the J_H band relative to embryo DNA, indicating extensive rearrangement of the J_H region. One sample (B1) shows no distinct J_H band at all, whereas the other duplicate samples (B2a, b) show very faint bands over an elevated background.

Because of the large difference in intensity of the J_H and C_μ bands it was not possible to select a range of autoradiographic exposures for which both bands were on the linear portions of their dose-response curves. Therefore, we had to use a different method of calculating the extent of J_H rearrangement from that used for the J_κ measurements. By varying the exposure time of the autoradiograph and scanning the tracks on a densitometer, we obtained calibration curves relating C_μ or J_H peak height to the input amount of embryo DNA (Fig. 5B). By reference to these curves we used the measured peak heights to estimate the apparent C_μ gene and J_H gene dosage as equivalent μ g total embryo DNA input (Table 1). The extent to which apparent J_H gene dosage is less than apparent C_μ gene dosage in B-cell DNA equals the extent of J_H rearrangement. Assuming that all IgM-expressing B cells must have one rearranged J_H region, we calculated the per cent of cells in each sample which have rearranged their silent J_H allele as: B1 ~100%, B2a 87% and B2b 81%. The persistence of a small proportion (~15%) of cells with an unrearranged J_H locus is consistent with the results of the plasmacytoma survey described above.

Discussion

These results clearly demonstrate that the aberrant rearrangements previously observed in transformed lymphoid cells also occur in normal B cells, and consequently that the recombination process responsible for the formation of functional immunoglobulin genes is error prone. Thus, when recombination events occur during B-lymphocyte maturation, a germ-line locus (Ig⁰) may be rearranged to form either a functional (Ig⁺) or a non-productive (Ig⁻) immunoglobulin gene. Detailed analyses of κ genes have revealed several types of non-productive rearrangements: those involving recombination at a pseudo J site, which lacks a proper RNA processing signal^{16,14,15}, those involving joining errors, which create translational stop codons^{16,17}, and those which extinguish transcription^{6,7}, possibly due to the selection of a non-functional (pseudo?) V-region segment.

We have suggested that the status of immunoglobulin alleles depends principally on their recombination frequency, which may vary over the course of B-cell maturation, and the relative probability of productive as opposed to non-productive rearrangements⁷. With appropriate values of these parameters, this probabilistic model should predict the ratio of Ig⁰/Ig⁺ to Ig⁻/Ig⁺ genotypes in a particular population of lymphoid cells. Our previous studies of κ -expressing plasmacytomas indicated a 1:1 ratio of κ^0/κ^+ to κ^-/κ^+ genotypes, whereas our current measurements of κ -expressing splenic B cells indicates a ratio of ~2:1. If this discrepancy actually reflects a difference between plasma cells and B cells, it would suggest that κ rearrangement can still occur during clonal expansion of B cells, although at a greatly diminished frequency compared with pre-B cells. Our estimate of 33% non-productive rearrangement in B cells is in marked contrast to the very low level claimed by Joho and Weissman⁸. However, when their single measurement was normalized with their rDNA data, we calculated 26% non-productive rearrangement, which is not significantly different from our more accurately determined value.

The high frequency of κ rearrangement in λ expressors and the rarity of λ rearrangement in κ expressors suggest that the recombination frequency is in part determined by the total

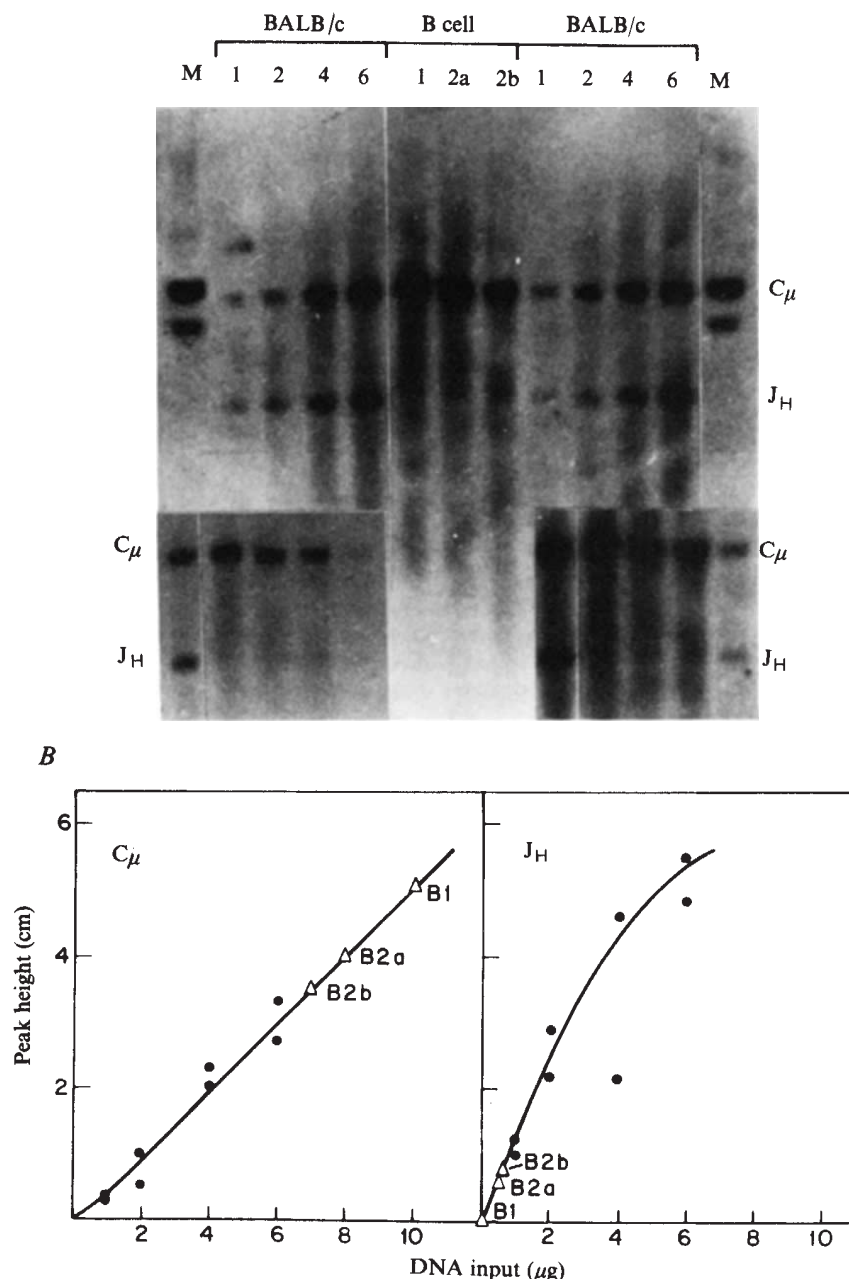
number of V and J elements that can participate in recombination, which in the mouse would greatly favour κ over λ . Furthermore, the fact that a κ^0 locus is occasionally retained by a λ expressor and a λ^- locus is occasionally present in a κ expressor is consistent with a probabilistic model which views the genes for both classes of light chain as equivalent candidates for productive or non-productive rearrangement. This fact is not easily accommodated by models which postulate a strictly sequential sampling of the κ and λ gene pools¹⁸.

If we consider the formation of a productive light-chain gene as the prime determinant of the developmental transition from pre-B to B cell¹⁹, then the composition of κ genotypes in λ -expressing spleen cells should reflect the steady-state distribution in pre-B cells at the time of this transition. Our results indicate a high proportion of κ^-/κ^- genotypes in these pre-B cells. For this to be compatible with a 2:1 ratio of κ^0/κ^+ to κ^-/κ^+ genotypes in B cells we need to postulate that the recombination frequency is greatly reduced after the formation of a productive light-chain locus. Such a supposition seems reasonable as the production of light chains allows the cell to display surface immunoglobulin²⁰ which presumably triggers

gross morphological and biochemical changes associated with the pre-B to B cell transition. This effective stabilization of genotypes by a productive rearrangement would effectively ensure the allelic exclusion of light-chain gene expression by diminishing the likelihood of $\kappa^0/\kappa^+ \rightarrow \kappa^+/\kappa^+$ transitions. A similar suggestion was recently made by Alt *et al.*¹⁸.

Our analysis of the heavy-chain (J_H) locus indicates a high frequency of non-productive rearrangement, with >80% of splenic B cells having μ^-/μ^+ genotypes. Yet the fact that a small but definite subpopulation have μ^0/μ^+ genotypes is consistent with the probabilistic nature of the recombination process. The difference in frequency of non-productive rearrangements between heavy- and light-chain genes may be related to the more complex structure of the heavy-chain locus, which requires at least two recombinational events (V_H to D_H and D_H to J_H) to produce a functional genetic unit. Lack of function may be due to an error in the D_H - J_H recombination, an error in the accompanying V_H - D_H fusion, or even the lack of a V_H - D_H recombination. In the latter case, the non-productive locus might represent an intermediate which could eventually undergo further rearrangement.

Fig. 5 J_H gene rearrangement in splenic B cells. **A**, B-cell-enriched fraction prepared as described in Fig. 1 legend was stained with fluorescein-coupled rabbit anti-mouse IgM and the surface IgM-positive fraction purified by FACS. The proportion of IgM-bearing cells in the final cell sample was estimated to be >98%. DNA was prepared from these cells and its concentration roughly estimated by visual inspection of ethidium bromide fluorescence of an aliquot compared with that of DNA of known concentration. B-cell DNA was digested sequentially with *SacI* and *EcoRI* and an estimated 6–8 μ g electrophoresed through 0.8% agarose. Two different samples of B-cell DNA were analysed, one of them in duplicate (B1; B2a, B2b). A titrated series of BALB/c embryo DNA was similarly analysed, in duplicate, loading 1, 2, 4 and 6 μ g as shown. Hybridization of the Southern blot to a mixture of ³²P-labelled plasmids p μ (3741)⁹ and pJ11 revealed the C_μ gene and the J_H region as fragments of 5.1- and 3.0-kilobases, respectively. Three exposures (20, 40 and 100 h) of the autoradiograph were developed. The entire 40-h exposure is shown, with the central five tracks (the three B-cell samples flanked on the left by 6 μ g and on the right by 1 μ g embryo DNA) of the 20- and 100-h autoradiographs shown as inserts on the lower left and lower right, respectively. **B**, the autoradiographs were scanned on a densitometer and peak heights measured. Peak heights of the C_μ fragment in embryo DNA from the 20-h autoradiograph were plotted against μ g input DNA to give a calibration curve linear at large DNA inputs, which we could use to estimate B-cell DNA input from C_μ peak height as shown. The curve relating embryo DNA input to J_H peak height on the 100-h autoradiograph is linear at small DNA inputs, and was used to estimate B-cell J_H gene dosage. ●, Embryo DNA; △, B-cell DNA.



A higher intrinsic recombinational frequency for heavy-chain compared with light-chain genes could explain why rearrangement of heavy-chain genes precedes that of light-chain genes during B-cell ontogeny^{21,22}. The fact that a relatively high proportion of the heavy-chain gene rearrangements are non-productive could account for the allelic exclusion of heavy-chain expression, for it would predict that most cells in which a $\mu^0 \rightarrow \mu^+$ transition occurs would already have experienced a $\mu^0 \rightarrow \mu^-$ on the other allele. This explanation is compatible with the idea that both heavy- and light-chain genes use the same recombinational machinery^{4,5}, and that this machinery remains essentially intact on formation of a functional heavy-chain gene. This seems reasonable because there are apparently no gross phenotypic changes in μ -expressing pre-B cells before the creation of a functional light-chain gene. Thus, allelic exclusion

of immunoglobulin gene expression is viewed as being due principally to recombinational errors in the case of heavy chains, and to a combination of errors and phenotypic alteration in the case of light chains.

We thank K. Marcu for the pJ11 plasmid, R. Riblet and J. Brill for many DNA samples, J. Haimovich for lymphoma DNA, B. Subbarao for fluorescent antibody and R. Overton for assistance with use of the FACS. The research was supported by a NSF grant, NIH institutional grants, a fellowship to K.K. from the EMBO and an appropriation from the Commonwealth of Pennsylvania.

Note added in proof: After submission of this article another observation of frequent rearrangement of the non-expressed J_H allele in normal B cells was reported³².

Received 16 September 1980; accepted 4 February 1981.

- Brack, C., Hiram, M., Lenhard-Schuller, R. & Tonegawa, S. *Cell* **15**, 1–14 (1978).
- Max, E., Seidman, J. G. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3450–3454 (1979).
- Sakano, H., Huppi, K., Heinrich, G. & Tonegawa, S. *Nature* **280**, 288–294 (1979).
- Early, P., Huang, H., Davis, M., Calame, K. & Hood, L. *Cell* **19**, 981–992 (1980).
- Sakano, H., Maki, R., Kurosawa, Y., Roeder, W. & Tonegawa, S. *Nature* **286**, 676–683 (1980).
- Perry, R. P. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 1937–1941 (1980).
- Perry, R. P., Coleclough, C. & Weigert, M. *Cold Spring Harb. Symp. quant. Biol.* (in the press).
- Joho, R. & Weissman, I. *Nature* **284**, 179–181 (1980).
- Azuma, T., Steiner, L. A. & Eisen, H. N. *Proc. natn. Acad. Sci. U.S.A.* **78**, 569–573 (1981).
- Karjalainen, K., Bang, B. & Mäkelä, O. *J. Immun.* **125**, 313–317 (1980).
- Coleclough, C., Cooper, D. & Perry, R. P. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1422–1426 (1980).
- Hurwitz, J. L., Coleclough, C. & Cebra, J. J. *Cell* **22**, 349–359 (1980).
- Marcu, K., Banerji, J., Pennacave, N. A., Lang, R. & Arnheim, N. *Cell* **22**, 187–196 (1980).
- Choi, E., Kuehl, M. & Wall, R. *Nature* **286**, 776–779 (1980).

- Seidman, J. G. & Leder, P. *Nature* **286**, 779–783 (1980).
- Altenburger, W., Steinmetz, M. & Zachau, H. G. *Nature* **287**, 603–607 (1980).
- Max, E. E., Seidman, J. G. & Leder, P. *Cell* **21**, 793–794 (1980).
- Alt, F. W., Enea, V., Bothwell, A. L. M. & Baltimore, D. *Cell* **21**, 1–12 (1980).
- Levitt, D. & Cooper, M. D. *Cell* **19**, 617–625 (1980).
- Paige, C. J., Kincade, P. W. & Ralph, P. J. *Immun.* **121**, 641–647 (1978).
- Maki, R., Kearney, J., Paige, C. & Tonegawa, S. *Science* **209**, 1366–1369 (1980).
- Perry, R. P., Kelley, D. E., Coleclough, C. & Kearney, J. *Proc. natn. Acad. Sci. U.S.A.* **78**, 247–251 (1981).
- Lenhard-Schuller, R., Hohn, B., Brack, C., Hiram, M. & Tonegawa, S. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4709–4713 (1978).
- Perry, R. P., Kelley, D. E., Schibler, U., Huebner, K. & Croce, C. M. *J. cell. Physiol.* **98**, 553–560 (1979).
- Southern, E. M. *J. molec. Biol.* **98**, 503–517 (1975).
- Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. *J. molec. Biol.* **113**, 237–251 (1977).
- Kearney, J., Radbruch, A., Liesgand, B. & Rajewsky, K. *J. Immun.* **123**, 1548–1550 (1979).
- Davis, M. M. *et al. Nature* **283**, 733–739 (1980).
- Wilson, R., Miller, J. & Storb, U. *Biochemistry* **18**, 5013–5021 (1979).
- Steinmetz, M., Zachau, H. G. & Mach, B. *Nucleic Acids Res.* **6**, 3213–3229 (1979).
- Cory, S., Adams, J. M. & Kemp, D. J. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4943–4947 (1980).
- Nottenburg, C. & Weissman, I. L. *Proc. natn. Acad. Sci. U.S.A.* **78**, 484–488 (1981).

LETTERS

Cyclotron and annihilation lines in γ -ray bursts

E. P. Mazets, S. V. Golenetskii, R. L. Aptekar', Yu. A. Gur'yan & V. N. Il'inski

A. F. Ioffe Physico-Technical Institute, Academy of Sciences of the USSR, 194021 Leningrad, USSR

Venera 11 and Venera 12 have detected emission and absorption lines in the energy spectra of many γ -ray bursts. The similarity between the spectral features observed during different events reveals the similarity in physical conditions and processes which occur within the burst sources. These lines are the most convincing evidence that the γ -burst sources contain neutron stars with a strong magnetic field. We present here results of a comprehensive analysis that included weaker bursts, which indicate that the presence of lines in the γ -burst spectra, in particular of absorption lines in the region 30–70 keV is a characteristic rather than an exceptional feature.

The investigation of lines in cosmic X-ray and γ -radiation is important, and has been reviewed elsewhere^{1,2}. γ -ray line radiation is characteristic of nuclear processes, however, in certain conditions, lines can be produced in the astrophysical plasma by radiative processes^{3,4}. Lines in the plasma radiation can be closely associated with the presence of superstrong gravitational and magnetic fields in sources such as degenerate dwarfs and neutron stars. The magnetic bremsstrahlung at the cyclotron frequency $\omega_H = eH/mc$ at fields $H > 10^{12}$ Oe already lies in the hard X-ray range. Pair production in the hot magnetized plasma may be a powerful source of positrons. The positron annihilation radiation produced near the surface of a neutron star should undergo a substantial gravitational redshift. Observations of these effects could provide a means for reliable determination of the gravitational and magnetic fields in the sources.

Cyclotron radiation in the X-ray range was first observed⁵ as the 58-keV line in the spectrum of Her X-1; the 73-keV line in the Crab spectrum may have the same nature⁶. It is thought that the redshifted annihilation line was observed⁷ among other lines in a transient event on 10 June 1974 and in the emission from the Crab nebula region⁸. However, not all of these pioneering

Table 1 Emission lines in the γ -ray burst spectra

Date of event	T_0 (UT)	Full energy flux S (erg cm ⁻²)	$\bar{E}$ (keV)	F (cm ⁻² s ⁻¹)	Energy in line S_e (erg cm ⁻²)	S_e/S
18 September 1978	19.49.11	2.8×10^{-5} (2.5×10^{-5})	400 ± 50	0.43 ± 0.10	$(2.1 \pm 0.5) \times 10^{-6}$	0.07 (0.08)
4 November 1978	16.17.53	2.6×10^{-4} (1.4×10^{-4})	400 ± 50	1.22 ± 0.19	$(5.7 \pm 0.9) \times 10^{-6}$	0.02 (0.04)
19 November 1978	09.28.33	5.5×10^{-4} (5.0×10^{-4})	330–850	15.0 ± 1.6	$(4.3 \pm 0.5) \times 10^{-5}$	0.08 (0.09)
16 January 1979	08.59.18	2.3×10^{-5}	420 ± 50	0.19 ± 0.07	$(1.9 \pm 0.7) \times 10^{-6}$	0.08
5 March 1979	15.51.39	1.3×10^{-3} (3.0×10^{-4})	430 ± 30	73 $\pm$ 15	$(1.0 \pm 0.2) \times 10^{-5}$	0.008 (0.03)
18 April 1979	07.44.41	6.5×10^{-5}	450 ± 40	0.26 ± 0.08	$(2.7 \pm 0.8) \times 10^{-6}$	0.04
26 May 1979	21.50.01	7.4×10^{-6}	45 $\pm$ 5	0.9 ± 0.1	$(7.9 \pm 0.8) \times 10^{-7}$	0.11
22 June 1979	00.41.30	7.2×10^{-5}	460 ± 40	0.12 ± 0.04	$(3.5 \pm 1.0) \times 10^{-6}$	0.05

Fig. 1 The energy spectra of γ -bursts. *a*, 7 March 1979; *b*, *c*, 1 November 1979 (illustrating the cyclotron absorption line evolution in time); the spectra shown were obtained in the first 8 s (*b*) and subsequent 24 s (*c*) of the burst; *d*, 12 June 1979; *e*, 12 April 1979 (initial stage); *f*, 10 June 1979.

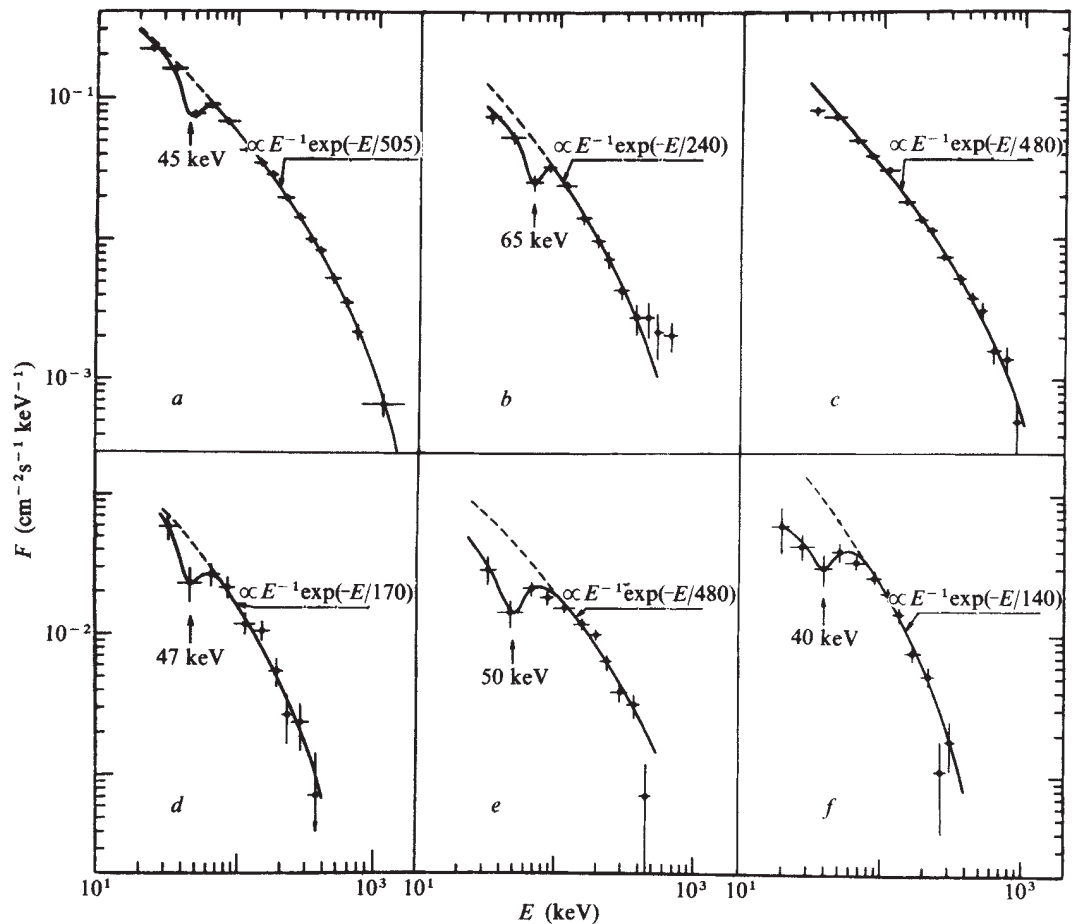


Table 2 Absorption lines in the γ -ray burst spectra

Date of event	T_0 (UT)	Full energy flux S (erg cm $^{-2}$)	$\bar{E}$ (keV)	Relative depth h	Equivalent width W (keV)	Fraction of energy absorbed S_a (erg cm $^{-2}$)	S_a/S
14 September 1978	16.42.12	1.7×10^{-5}	65 ± 10	Wide absorption band	28 ± 5	$(6.2 \pm 1.0) \times 10^{-7}$	0.05 (0.08)
17 November 1978	13.46.42	1.2×10^{-5} (8.0×10^{-6})	52 ± 8				
21 November 1978	01.34.14	9.0×10^{-5} (1.5×10^{-5})	70 ± 10				
7 March 1979	22.15.34	1.9×10^{-4}	45 ± 5	0.43	7.0 ± 0.6	$(2.3 \pm 0.2) \times 10^{-6}$	0.012
23 March 1979	23.39.11	3.4×10^{-5} (1.1×10^{-5})	66 ± 10	0.43	30 ± 5	$(1.1 \pm 0.2) \times 10^{-6}$	0.03 (0.10)
25 March 1979	13.41.12	4.8×10^{-5} (1.9×10^{-5})	50 ± 10	0.32	12.0 ± 1.3	$(5.6 \pm 0.6) \times 10^{-7}$	0.012 (0.03)
29 March 1979	22.28.42	6.5×10^{-5} (3.7×10^{-5})	50 ± 8	0.35	3.9 ± 0.7	$(2.6 \pm 0.5) \times 10^{-7}$	0.004 (0.007)
2 April 1979	09.41.45	3.7×10^{-5}	55 ± 15	Wide absorption band	27 ± 2	$(1.8 \pm 0.2) \times 10^{-6}$	0.08 (0.10)
12 April 1979	22.04.42	2.2×10^{-5} (1.7×10^{-5})	50 ± 8				
24 May 1979	03.56.26	1.5×10^{-5}	27 ± 5	0.31	3.4 ± 0.9	$(1.9 \pm 0.5) \times 10^{-7}$	0.013
29 May 1979	20.23.01	7.4×10^{-6}	55 ± 10	0.46	20 ± 5	$(3.5 \pm 0.9) \times 10^{-7}$	0.05
10 June 1979	05.50.04	6.1×10^{-6}	40 ± 5	Wide absorption band	13 ± 4	$(2.5 \pm 0.8) \times 10^{-7}$	0.08
12 June 1979	04.42.54	3.0×10^{-6}	47 ± 7				
22 June 1979	00.41.30	7.2×10^{-5}	50 ± 8				
31 July 1979	10.49.55	8.1×10^{-5} (3.1×10^{-5})	55 ± 10				
6 October 1979	08.10.26	7.3×10^{-6} (2.2×10^{-6})	62 ± 8	0.52	28 ± 4	$(4.0 \pm 0.6) \times 10^{-7}$	0.055 (0.18)
31 October 1979	09.18.33	6.5×10^{-6} (5.4×10^{-6})	46 ± 8	Wide absorption band	18 ± 2	$(5.9 \pm 0.8) \times 10^{-7}$	0.003 (0.017)
1 November 1979	17.36.27	2.1×10^{-4} (3.4×10^{-5})	65 ± 10				
9 November 1979	07.56.32	6.8×10^{-6}	64 ± 10				
20 December 1979	17.30.34	1.0×10^{-4} (5.7×10^{-5})	63 ± 10	0.22	18 ± 1	$(1.6 \pm 0.1) \times 10^{-6}$	0.016 (0.028)

results are reliable²; in particular, the spectral features from the Crab nebula have not shown up in other observations therefore independent confirmation is needed.

Energy spectrum studies and the search for lines are also important in understanding the nature of the γ -bursts. In the Konus experiment carried out on Venera 11 and Venera 12 spacecraft during 1978–80 (ref. 9) energy spectra of ~ 150 γ -ray bursts in the 30 keV–2 MeV range were obtained, and the very first observations¹⁰ revealed lines in the spectra of some bursts. A redshifted annihilation line was clearly observed at 430 keV in the spectrum of the 5 March 1979 burst produced by a flaring X-ray pulsar¹¹. A broad emission feature was detected at 300–800 keV in the spectrum of a strong γ -burst on 19 November 1978 (ref. 9). A germanium detector on ISEE 3 revealed this as two lines at ~ 420 and 740 keV which were interpreted as the strongly redshifted 511-keV annihilation line and 847-keV iron line, respectively¹². Our subsequent measurements showed the presence of emission and absorption lines in several strong γ -bursts¹³.

The most typical examples of γ -burst spectra containing lines are shown in Figs 1–3. (More information on these events is given in refs 14, 15.) Eight consecutive measurements of the energy spectrum were made during each event, each lasting 4 s. Thus between 1 and 8 spectra were obtained for each burst, depending on its duration. The statistical accuracy of an individual spectrum of a weak event may be insufficient for line resolution, in which case, individual spectra are added together. This summation can also be used if no noticeable evolution of the spectrum is observed in the course of a burst.

Figure 1a shows the spectrum of a long intense event on 7 March 1979. Here the spectral distribution in the continuum, just as in most of the other bursts, fits the $dN/dE \propto$

$E^{-1} \exp(-E/kT)$ law typical for radiation from a hot, optically-thin plasma. An absorption line can be clearly seen at ~ 45 keV remaining visible in all the measured spectra. The line is assumed to result from absorption of the burst radiation at the cyclotron frequency ω_H in the outer, colder regions of the plasma in a strong magnetic field. The resonance is associated with the change of state of the electron in a magnetic field, its transition to the next Landau level involving the absorption of an energy $\Delta E = \hbar\omega_H = \hbar eH/mc$.

The spectrum of the 1 November 1979 burst is shown in Fig. 1b, c. In contrast to the 7 March 1979 event, the absorption line is observed here only at the initial burst stage (Fig. 1b). At the later stage (Fig. 1c) the line disappears—an evolution typical of many other bursts which may be associated with the heating of the absorption region during the burst development. The increase in emission at the cyclotron frequency in this region due to heating can cancel absorption in the spectrum of the γ -burst. In the 1 November 1979 burst the spectral distribution in the continuum also undergoes evolution. The spectrum becomes harder and the temperature kT increases from 240 to 480 keV.

An interesting feature of the absorption line in the spectrum of the 1 November 1979 event is an intensive wing at lower energies which causes a marked broadening of the equivalent line width. Apparently, this can be attributed to a spatial variation in magnetic field in the source. Such an interpretation also holds for many other bursts, for example, those with the spectra shown in Fig. 1d–f. Here the absorption line spreads into a band because of a substantial change in magnetic field within the source.

Figure 2a, b shows spectra with absorption lines obtained for the bursts on 29 March 1979 and 24 May 1979. Note that these bursts were recorded by the Venera 11 detector whose energy

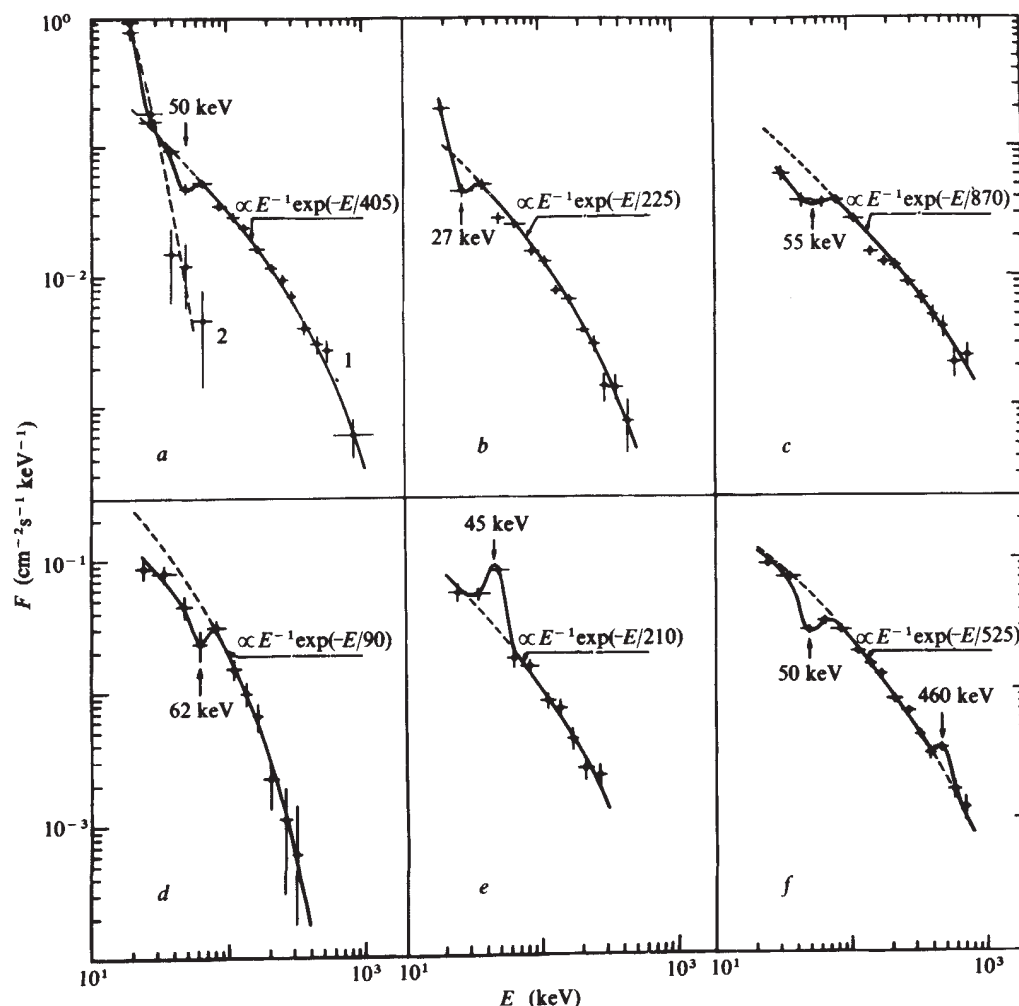
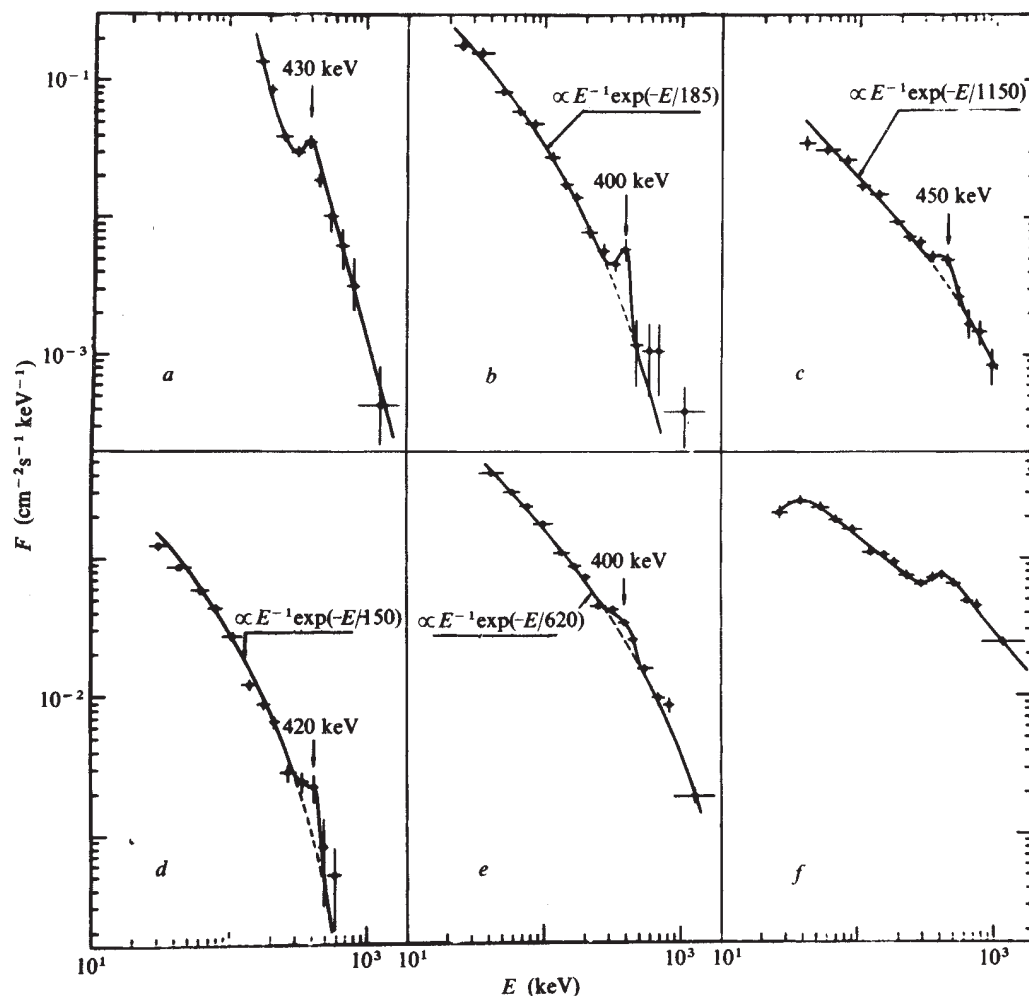


Fig. 2 The energy spectra of γ -bursts. *a*, 29 March 1979 (1, the spectrum of the initial stage of the burst; 2, the spectrum obtained for the deep minimum between the burst peaks); *b*, 24 May 1979; *c*, 31 July 1979 (initial stage); *d*, 6 October 1979 (initial stage); *e*, 26 May 1979; *f* 22 June 1979.

Fig. 3 The energy spectra of γ -bursts with emission lines. *a*, 5 March 1979 (the spectrum of the short initial pulse); *b*, 18 September 1978; *c*, 18 April 1979; *d*, 16 January 1979; *e*, 4 November 1978 (initial stage); *f*, 19 November 1978 (initial stage).



threshold lies at ~ 17 keV in contrast to the other five detectors with a threshold of ~ 30 keV. This enabled us to obtain additional information. The radiation level in the first, the softest channel of the spectra of these bursts, is higher than normal and remains nearly constant during the burst. The 29 March 1979 event consists of many peaks and lasts about 40 s (ref. 14). The last measured spectrum (2 in Fig. 2a) falls on the deep minimum between the peaks—its shape corresponds to radiation from a substantially colder plasma with temperature $kT \approx 8$ keV. A comparison of the two spectra in Fig. 2a shows that the soft component of the radiation is present during the entire burst. This suggests that the bursts on 29 March 1979 and 24 May 1979 come from two-component sources. It is unclear whether the appearance of the absorption lines can be associated with the presence in the source of a substantially more stationary and colder plasma.

Figure 2c, d shows absorption at the initial stages of the 31 July 1979 and 6 October 1979 bursts which also undergoes evolution, disappearing at later stages. All these graphs demonstrate the similarity between the patterns of absorption in the spectra with substantially different spectral distributions in the continuum.

The burst on 26 May 1979 (Fig. 2e) revealed an emission line at 45 keV. From our data, this spectrum is the only reliable evidence of emission at the cyclotron frequency ω_H . Figure 2f shows a spectrum of the 22 June 1979 burst which is remarkable in that it contains in addition to the cyclotron absorption line and an emission line at 460 keV which is apparently due to the annihilation radiation. The energy of this line was reduced $(1 - 2GM/c^2R)^{1/2}$ times as a result of gravitational redshift in the field of a neutron star.

The same emission lines were also observed in several other γ -ray bursts (Fig. 3). These data show that the relative intensity

of the emission lines, just as that of the absorption lines, can vary in a time shorter than the burst duration. The line emission seems most likely to appear at the burst stages characterized by the highest temperature of radiation in the continuum. Indeed, the line at ~ 430 keV in the burst on 5 March 1979 can be seen only in the hard spectrum of the short initial pulse¹¹. A broad line at ~ 400 keV simultaneous with hard radiation in the continuum can only be observed at the initial stage of the 4 November 1978 burst. The event on 19 November 1978 is distinctive because of its extremely strong evolution of the spectral distribution⁹. The very hard spectrum with a broad emission feature at 300–800 keV (Fig. 3f) is observed during the first 4 s of the burst. The resolution in this region of two lines at 420 and 740 keV (ref. 12) apparently represents the only case of observation in a γ -ray burst of a line above 500 keV. The temporal evolution of the emission lines in the 18 September 1978, 18 April 1979, and 16 January 1979 bursts is less pronounced.

Tables 1 and 2 contain data on lines in the γ -ray bursts, showing the date of the event, the time of the beginning of recording T_0 , and the energy flux S (>30 keV) integrated over the burst duration. For the emission lines the mean energy $\bar{E}$, the mean photon flux F , the energy flux S_e emitted in the line, and the ratio S_e/S are given. The parameters for the absorption lines are the mean energy $\bar{E}$, the relative line depth h in units of the continuum intensity, the estimated equivalent line width $W = \int h dE$, the estimated fraction of absorbed energy flux in the burst S_a , and the ratio S_a/S . If evolution of a line was observed, the values of S and S_e/S , or S_a/S calculated for the burst stage with the line are also given in brackets.

The figures and tables show that the lines in different bursts have fairly similar characteristics. Emission lines appear predominantly in the interval 400–460 keV and the energy

contained in these constitute 5–10% of the energy of the burst. Absorption lines are observed in the 30–70-keV region and the equivalent line widths constitute 5–15 keV, thus 1–2% of the energy of the burst is absorbed. In some cases the lines become smeared strongly and the integral absorption in the soft region of the spectrum increases. The temporal evolution of lines in different bursts also reveals common features. Such a similarity supports our interpretation of the spectral features in the γ -burst spectra as cyclotron absorption lines and redshifted annihilation lines. Thus an overall analysis of our data leads to a conclusion that γ -bursts are produced in regions with a gravitational potential of $\sim(0.1\text{--}0.2)c^2$ and magnetic field $\sim(2\text{--}6) \times 10^{12}$ Oe near the surface of neutron stars. Strong support for these neutron stars being actual components of binaries¹⁶ is the discovery of a similar cyclotron absorption line at 20 keV in the spectrum of the pulsed hard X-ray flux¹⁷ from 4U0115+63.

Received 28 October; accepted 21 December 1980.

1. Ramaty, R. & Lingenfelter, R. E. *Nature* **278**, 127–132 (1979).
2. Vedrenne, G. *Adv. Space Explor.* **7**, 85–99 (1979).
3. Gnedin, Yu. N. & Sunyaev, R. A. *Astron. Astrophys.* **36**, 379–394 (1974).
4. Canuto, V. & Chiu, H. Y. *Space Sci. Rev.* **12**, 3–74 (1971).
5. Trümper, J. *et al. Astrophys. J. Lett.* **219**, L105–L110 (1978).
6. Ling, J. C. *et al. Astrophys. J.* **231**, 896–905 (1979).
7. Jacobson, A. S. *et al. Gamma Ray Spectroscopy in Astrophysics* (eds Cline, T. L. & Ramaty, R.) 228–251 (NASA TM 79619, 1978).
8. Leventhal, M., MacCallum, J. C. & Watts, A. C. *Astrophys. J.* **216**, 491–502 (1977).
9. Mazets, E. P. *et al. Kosm. Issled.* **17**, 812–819 (1979).
10. Mazets, E. P. *et al. Preprint no. 599* (A. F. Ioffe Physico-Technical Institute, 1979).
11. Mazets, E. P. *et al. Pisma Astr. Zh.* **5**, 307–312 (1979); *Nature* **282**, 587–589 (1979).
12. Teegarden, B. J. & Cline, T. L. *Astrophys. J. Lett.* **236**, L67–L70 (1980).
13. Mazets, E. P. *et al. Pisma Astr. Zh.* **6**, 706–711 (1980).
14. Mazets, E. P. *et al. Preprints nos 618 & 637* (A. F. Ioffe Physico-Technical Institute 1979).
15. Mazets, E. P. *et al. Preprint no. 662* (A. F. Ioffe Physico-Technical Institute, 1980); *Astrophys. Space Sci.* (in the press).
16. Mazets, E. P. & Golenetskii, S. V. *Preprint No. 632* (A. F. Ioffe Physico-Technical Institute, 1979); *Astrophys. Space Sci.* (in the press).
17. Wheaton, Wm. A. *et al. Nature* **282**, 240–243 (1979).

Observations of a circumstellar shell around the OH/IR star OH127.8–0.0

R. S. Booth, A. J. Kus*, R. P. Norris & N. D. Porter

University of Manchester, Nuffield Radio Astronomy Laboratories, Jodrell Bank, Macclesfield, Cheshire SK11 9DL, UK

The 1,612 MHz OH maser emission associated with cool long-period variable stars is generally characterized by two main regions of emission separated in velocity by $\sim 10\text{--}50 \text{ km s}^{-1}$. This has been taken to indicate that the OH lies in an expanding shell surrounding the star, the main peaks originating at the near or far sides of the shell along the line of sight where the maser column lengths may be greatest. Here we present the first high-angular resolution measurements of the OH emission from such a source which demonstrates dramatically that this picture is correct.

The 1,612-MHz OH maser source, OH127.8–0.0 (hereinafter called OH127.8) was discovered in a survey by Bowers¹. Like many other sources found in such surveys^{2,3}, it shows the OH characteristics of the long period variable stars but has no optical identification although it is probably coincident with an IR star⁴. These objects have been designated Type II OH/IR stars and are assumed to be long period variable stars which are either at too large a distance for optical identifications or have circumstellar shells which are too thick for the visible radiation to penetrate (their properties are discussed elsewhere^{5,6}).

Observations of the OH maser radiation from OH127.8 were made between late 1978 and early 1980 with three single baseline interferometers of the Jodrell Bank telescope network, before the completion of the new Multi Telescope Radio Linked

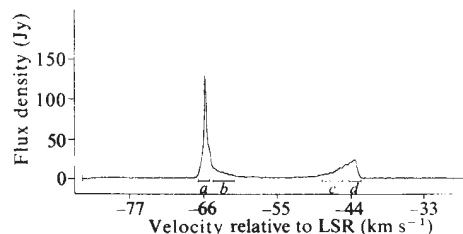


Fig. 1 The 1,612 MHz OH spectrum of OH127.8. The bracketed velocity intervals *a*, *b*, *c* and *d* represent those ranges of velocity for which the spatial distribution of the maser emission has been determined (see Fig. 2).

Interferometer (MTRLI)⁷ and as a prelude to its use as a line instrument.

These observations of OH maser radiation form part of a larger programme of observations of several such objects using the MkIA–Knockin, the MkIA–Defford and the MkIA–MkIII interferometers respectively⁸. On each occasion the usual radio link technique was used and the data were cross-correlated in the 1,024-channel digital spectrometer to give a 512 point cross-correlation spectrum in the manner described by Norris and Booth¹⁸.

The autocorrelation spectrum of OH127.8 is shown in Fig. 1. The special feature of this source is that the narrow spectral component at -65.7 km s^{-1} was virtually unresolved on all baselines and could therefore be used as phase reference not only to correct the phase on all the other channels of the cross-correlation spectrum but also to combine the three sets of interferometer data. After combination and normalization the new data set was split into velocity intervals and a series of maps produced, each corresponding to a particular range of velocities in the masing gas. The procedure, described fully elsewhere⁸ consisted essentially of Fourier inversion and cleaning⁹.

Four of the maps are reproduced in Fig. 2*a–d*—they represent the spatial distribution of the masing gas in the velocity intervals *a*, *b*, *c* and *d* indicated in Fig. 1. Figure 2*a* is a map of the distribution of the OH gas with velocities close to the reference feature and in the range -65 to -66.5 km s^{-1} (region *a* in Fig. 1). It shows that the material within this velocity interval is concentrated in a small region of angular extent $\sim 0.4 \text{ arc s}$ and is

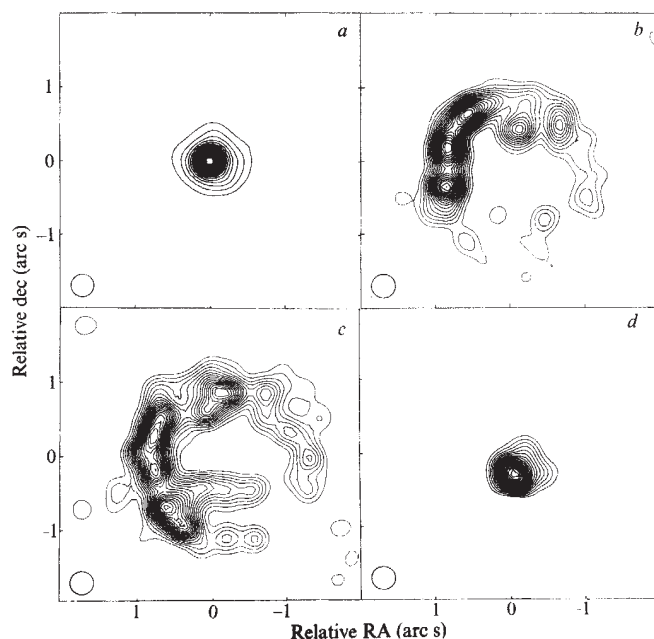


Fig. 2 *a–d*, Maps of the spatial distribution of the integrated OH emission in the velocity intervals marked *a* to *d* respectively in Fig. 1. In each map the contour interval is 5% of the peak integrated emission in that velocity interval and the lowest contour is $\sim 10\%$.

* Permanent address: Nicolaus Copernicus University, Gagarina 11, 87-100 Toruń, Poland.

dominated by the -65.7 km s^{-1} component. Figure 2d shows the distribution of OH in the corresponding redshifted velocity interval (-43 to -43.7 km s^{-1}) and is clearly similar in extent although slightly offset in position.

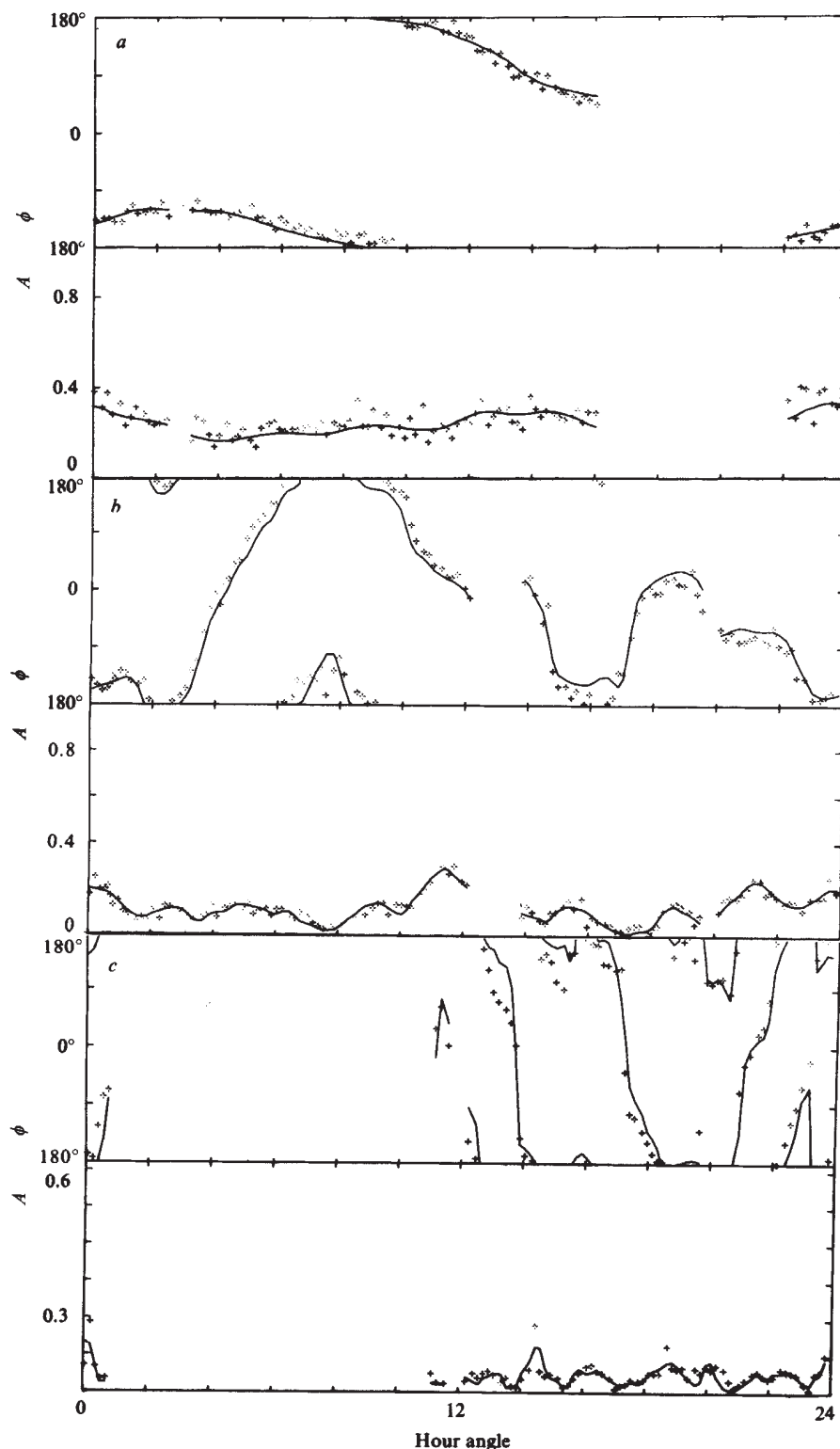
Figure 2b, c shows maps of the weaker OH emission in velocity regions closer to the stellar velocity (-62.2 to -64.7 km s^{-1} and -45.2 to -47.5 km s^{-1} respectively). Both have a similar annular form with angular diameter ~ 2 arc s. The map centres are the same in each case and correspond to the position of the reference feature.

The distributions found are exactly those expected of a uniformly expanding circumstellar shell¹⁰ of OH. The strongest

maser components have small angular diameters and are situated close to the line of sight through the centre of the shell along which the gain paths are largest, and the weaker emission delineates those portions of the shell where constant velocity path lengths along the line of sight will be shorter, giving lower maser gain. At the edges of the shell the velocity gradient is such that the line of sight emission is reduced to zero.

Note that the spectrum of the source varied slightly in intensity between observations thus we could not be sure of the relative amplitudes to better than 10%. However, as the fits to the data were good, we believe that our normalization was correct; as an example, in Fig. 3 we show a plot of the data and

Fig. 3 A plot of the visibility phase (ϕ) and amplitude (A) of the OH emission in the velocity range b of Fig. 1 together with the same data (solid line) derived from Fourier inversion of the map plotted in Fig. 2b. Base-lines, *a*, MkIA-MkII; *b*, MkIA-Knockin; *c*, MkIA-Deffprd.



corresponding points produced by Fourier inverting the map of Fig. 2b. As an extra check on this effect, we varied the amplitudes by 10% on each baseline and repeated the fitting procedures. In all cases, although the detailed distribution changed, the overall shell pattern remained, it being strongly dependent on the phase data which were always referenced to the unresolved point component at -65.7 km s^{-1} .

Assuming the uniformly expanding shell model¹⁰, the radius at which the line of sight component of the expansion velocity is v , $a(v)$ is given by

$$a(v) = R[1 - (v/v_e)^2]^{1/2}$$

where v_e is the expansion velocity and R the total shell radius. In the case of OH127.8, the expansion velocity (half the interval between OH peaks in Fig. 1) is $\sim 11 \text{ km s}^{-1}$. The rings of emission in Fig. 2b, c have mean velocity relative to the star of 8.5 km s^{-1} and angular diameter $\sim 2 \text{ arc s}$. Thus the total shell angular diameter is $\sim 3.1 \text{ arc s}$ which corresponds to a linear diameter of $1.5 \times 10^{17} \text{ cm}$ if we take the kinematic distance to the star to be 3.3 kpc^{11} .

A technique used recently to determine the linear diameter of the OH shells surrounding long period variable stars¹² involves careful observation of the OH variability when a phase difference can be observed in the light curve of the blueshifted emission from the nearside of the shell, relative to the redshifted emission from the far side. Assuming that the central star is the source of pump radiation, the OH emission from the entire shell should vary in unison and the observed phase lag represents the shell crossing time. Phase lag measurements by Jewell *et al.*¹³ and Herman and Habing¹⁴ show that diameters, d , lie in the range $10^{16} \text{ cm} < d < 10^{18} \text{ cm}$. Although these groups have not sufficiently monitored OH127.8 to obtain a definite phase lag result on this source, our value of diameter lies well within the range of measured values.

The mass of the circumstellar shell and particularly the mass loss rate from long period variable stars are of great importance in understanding stellar evolution and the chemical composition of the interstellar gas. Observations of the circumstellar masers will provide information on the masses involved if the gas density in the shell can be determined. This involves a determination of the OH column density through consideration of the brightness temperature of the maser, its degree of saturation and the efficiency of its pump mechanism, and an estimate of the abundance of OH relative to H_2 . In the case of OH127.8, individual maser features are unresolved with our longest baseline and they may be as small as 30 m arc s (ref. 11). Thus their brightness temperatures are $\sim 10^9 \text{ K}$. An efficient $35\text{-}\mu\text{m}$ IR pumping scheme for the masers in circumstellar envelopes has been developed by Elizur *et al.*¹⁵ and their calculations show that the $1,612 \text{ MHz}$ transition is inverted and saturated for OH densities $n_{\text{OH}} \geq 1 \text{ cm}^{-3}$. Brightness temperatures of $\sim 10^9$ are produced for OH column densities $\sim 10^{17} \text{ cm}^{-2}$, which agrees with our observations. A detailed model of the circumstellar OH abundance has been developed by Goldreich and Scoville¹⁶ who argue that $n_{\text{OH}} \geq 1 \text{ cm}^{-3}$, at a radius of $\sim 10^{16} \text{ cm}$ when the molecular hydrogen number density is $\sim 2 \times 10^4 \text{ cm}^{-3}$. This value, together with our measured linear diameter, gives a value of $\sim 10^{-1} M_{\odot}$ for the mass of the shell and a mass loss rate $\sim 8 \times 10^{-5} M_{\odot} \text{ yr}^{-1}$. Although such a mass loss rate is extremely high, it agrees well with rates derived for a sample of OH/IR stars from measurements of their IR properties, luminosity and dust shell opacities together with the outflow velocities¹⁷. Clearly, stars losing mass at these rates must be short-lived phenomena.

Finally OH/IR stars generally have angular sizes which lie in the resolution range of the MTRLI and observations are already being made using this instrument. As well as measuring the detailed structure and mass loss rates of these fascinating objects, a combination of the angular size of their shells from interferometry and the direct (phase lag) measurement of their diameter will lead to valuable independent estimates of their distances.

We acknowledge the assistance of Dr B. Anderson with some of the equipment and Drs P. N. Wilkinson, R. G. Noble and T. J. Cornwell for their advice on the mapping procedures. A.J.K. received a British Council support grant.

Received 17 December 1980; accepted 9 February 1981.

1. Bowers, P. F. *Astr. Astrophys. Suppl.* **31**, 127-145 (1978).
2. Johanson, L. E. B. *et al. Astr. Astrophys.* **54**, 323 (1977).
3. Baud, B. *et al. Astr. Astrophys.* (in the press).
4. Engels, D. *Astr. Astrophys. Suppl.* **36**, 337 (1979).
5. Snyder, L. E. in *Interstellar Molecules* (ed. Andrew, B. H.) (Reidel, Dordrecht, 1980).
6. Elitzur, M. in *Proc. Workshop on phys. Processes in Red Giants, Erice 1980* (eds Renzini, A. & Iben, I.) (in the press).
7. Davies, J. G., Anderson, B. & Morison, I. *Nature* **288**, 64 (1980).
8. Porter, N. D. *et al.* (in preparation).
9. Hogbom, J. A. *Astr. Astrophys. Suppl.* **15**, 417 (1974).
10. Reid, M. J. *et al. Astrophys. J.* **214**, 60 (1977).
11. Bowers, P. F. *et al. Astrophys. J.* **242**, 1088 (1980).
12. Schultz, G. V., Sherwood, W. A. & Winnberg, A. *Astr. Astrophys.* **63**, L5 (1978).
13. Jewell, P. R., Webber, J. C. & Snyder, L. E. *Astrophys. J.* **242**, L29 (1980).
14. Herman, J. & Habing, H. in *Proc. Workshop on phys. Processes in Red Giants, Erice 1980* (eds Renzini, A. & Iben, I.) (in the press).
15. Elizur, M., Goldreich, P. & Scoville, N. *Astrophys. J.* **205**, 384 (1976).
16. Goldreich, P. & Scoville, N. *Astrophys. J.* **205**, 144 (1976).
17. Werner, M. W. *et al. Astrophys. J.* **239**, 540 (1980).
18. Norris, R. P. & Booth, R. S. *Mon. Not. R. astr. Soc.* (in the press).

Fragmentation of a meteor in near-Earth space

M. A. Hapgood & P. Rothwell

Department of Physics, The University, Southampton SO9 5NH, UK

Fragments of a Perseid meteor, which broke up well above the Earth's atmosphere, probably in collision with a dust particle in near Earth space, have been observed with a two-station low-light level television camera system. Three Perseid meteoroids, with apparent magnitudes between 2.0 and 3.5, arrived on near parallel trajectories within 1.3 s of each other on the morning of 12 August 1977 (03.05.21-03.05.23 UT). The probability of these meteoroids arriving by chance so close together in time and direction is very small ($\sim 3 \times 10^{-11}$). Assuming the three meteoroids were fragments of a larger body, we show here that the parent meteor must have broken up at least 1,700 km above the Earth's surface, well above the altitude at which meteoroids normally start to burn up in the Earth's upper atmosphere (100-150 km). An estimate of the size of the parent meteor has been made from the observed magnitudes of these fragments. Assuming the parent meteoroid was spinning with an angular velocity between 100 and 1,000 rad s^{-1} , an upper limit to the distance of breakup point from the Earth's surface can be placed at between 7×10^5 and $7 \times 10^6 \text{ km}$ (0.004-0.04 AU). We show here that the probabilities of collision and breakup of the parent Perseid with three classes of dust particle in near Earth space—interplanetary, magnetospheric and lunar ejecta—at the time of our observation in 1977, were $\sim 3 \times 10^{-6}$, 3×10^{-6} and 5×10^{-7} respectively. We suggest therefore that if the Perseid parent comet (Swift-Tuttle) passes close to the Earth when it returns to perihelion in the next year or two, giving rise to a spectacular display of the Perseids, we should then be able to see many of these associated meteor events, and to study meteor collision with dust particles in near Earth space.

Table 1 Meteor parameters

	Meteor 1	Meteor 2	Meteor 3
Geocentric radiant:			
RA (α)	3 h 26 min	3 h 31 min	3 h 19 min
Dec (δ)	+53.7°	+54.2°	+53.8°
Geocentric velocity (km s^{-1})	59 $\pm$ 2	59 $\pm$ 2	56 $\pm$ 3
Apparent magnitude	2.0 $\pm$ 0.5	2.1 $\pm$ 0.5	3.5 $\pm$ 0.5
Magnitude at zenith	0.5	0.6	1.8
Mass (10^{-5} kg)	13	12	4

The meteor observations were made with low-light level TV cameras and associated video recording equipment at two sites in the south of England; Southampton (50°50' N, 1°24' W) and Harwell (51°35' N, 1°20' W). An image isocron camera, with a 30°×23° field of view, was used at Southampton and a silicon intensified target (SIT) camera, with a 20°×15° field of view, was used at Harwell. Both cameras looked west at a common volume centred above South Wales at a height of ~100 km. Fifty pictures per second were recorded on video tape; these pictures could be played back slowly frame by frame. Typically a meteor appeared in ~10 successive frames.

The positions in space of the three associated Perseids were determined by triangulation against the fixed star background using observations from the two sites. Figure 1 shows the variation of the altitude of the lower tip of the meteor trails with time; Fig. 2 shows the positions of the three trails projected onto the ground. The velocities and radiant of the three meteors are listed in Table 1; they all lie close to the mean Perseid values¹ (geocentric velocity 59.4 km s⁻¹, geocentric radiant $\alpha = 3$ h 05 min, $\delta = 57.4^\circ$). The error in the determination of each velocity is ~2 km s⁻¹; the error in the position of one radiant relative to another is ~1°. The apparent magnitude of each meteor was determined by comparing its brightness with the brightness of stars in the field of view. The meteor magnitudes, corrected to the zenith², are also listed in Table 1, together with the masses, estimated from the zenith magnitudes, following Hughes³.

The three meteors arrived within a time interval of 1.3 s and on trajectories which were parallel to within 1°, that is, within a solid angle of ~10⁻³ sr. In contrast the average meteor rate in the field of view of our cameras during the Perseid meteor shower was ~6 h⁻¹; the radiant of these meteors were spread over a wide region of the sky ~12° across (solid angle ~0.15 sr). Thus the probability of observing by chance three meteors arriving so close together in time and direction was ~3×10⁻¹¹.

By combining the masses of the three meteors (see Table 1) we can estimate that the mass of the parent body was ~3×10⁻⁴ kg. If we assume that the parent body was roughly spherical and had a typical Perseid density⁴ of 290 kg m⁻³ its radius was ~6 mm.

We can estimate the distance s from the Earth back to the breakup point since $s = v^2 \Delta t / \Delta v$, where v is the mean velocity of the fragments, Δv is their velocity dispersion and Δt is the time interval between the arrival of the fragments. The mean velocity ($v = 58.5 \pm 1.3$ km s⁻¹) and the time interval ($\Delta t = 1.30 \pm 0.02$ s, see Fig. 1) are known quite accurately. However, the velocity dispersion Δv lies within the errors of measurement so the error in the mean velocity sets an upper limit on Δv . We take $\Delta v_{\max}$ to be 2.6 km s⁻¹, which gives a value of the minimum distance to breakup point, $s_{\min}$ of 1,710 km. We can make a second and independent estimate of $s_{\min}$ from the observation that the

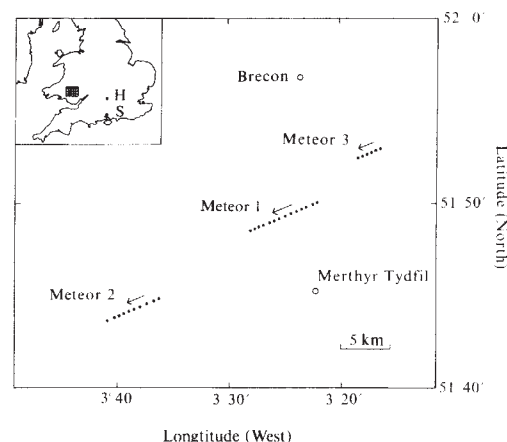


Fig. 2 The horizontal motion of the meteors. The position of each meteor is shown at time intervals of 20 ms (1 TV picture). The inset shows the area (shaded) of the main map and the positions of the two observing sites: Southampton (S) and Harwell (H).

meteor fragments travelled on paths which were parallel to within the errors of measurement (1°) and were ~30 km apart at 105-km altitude. Taking the maximum angle of dispersion between the trajectories to be 1° we obtain $s_{\min} = 1,720$ km which agrees well with the value deduced from the velocity dispersion. This puts the fragmentation point outside the Earth's atmosphere, thus a high velocity collision between the parent body and a small dust particle would be the likely cause of fragmentation.

We can set an upper limit to the distance of the breakup point from the Earth if we can set a minimum value for the velocity dispersion of the fragments ($\Delta v_{\min}$). In general we would expect a significant velocity dispersion to be generated by the collision. However, in the case of a collision with a small dust particle which is only just large enough to cause fragmentation, almost all the kinetic energy of the collision would be dissipated in breaking up the parent meteoroid. In such cases the only contribution to the velocity dispersion is that arising from the rotation of the parent body. Thus we may put $\Delta v_{\min} \sim \omega r$ where ω is the angular velocity of the parent body and r is its radius. Theoretical studies indicate that meteoroid spin may be produced by collisions⁵ and by solar radiation effects⁶. These theories indicate that angular velocities of 100–1,000 rad s⁻¹ may be achieved within the lifetime of the Perseid stream (~3,000 yr (ref. 7)). This implies a value for $\Delta v_{\min}$ in the range 0.6–6 m s⁻¹ giving a value $s_{\max}$ in the range 7–70×10⁵ km (0.004–0.04 AU). Thus we may conclude that the parent body broke up in space close to the Earth.

The minimum mass of dust particle needed to fragment the parent body can be estimated from the results of the hyper-velocity impact experiments of Moore and Gault⁸. They found that a body of mass M would be fragmented by the impact of a particle of mass m and relative velocity v km s⁻¹ when $m > M/250 v^2$. If we apply this formula to the present case we find that a dust particle of minimum mass 3×10⁻¹⁰ kg would be required to give fragmentation. However, note that the hyper-velocity impact experiments from which the formula was derived were carried out on basalt targets with relative velocities up to 7 km s⁻¹. The use of the Moore and Gault⁸ formula involves considerable extrapolation in velocity and assumes that the meteoroid material fragments in similar conditions to basalt.

We may estimate the probability of observing a group of associated meteors by calculating the probability of the parent body suffering a collision. Dohnanyi⁹ has calculated the collisional lifetimes T_c , of meteoroids in several meteor streams, with interplanetary dust particles, taking into account both the meteoroid cross-section and the number density of dust particles which are, according to Moore and Gault⁸, large enough to cause fragmentation. He finds that, for the Perseids, $T_c = 1.8 \times 10^4 M^{1/3}$ yr where M is the meteoroid mass in kg. Thus

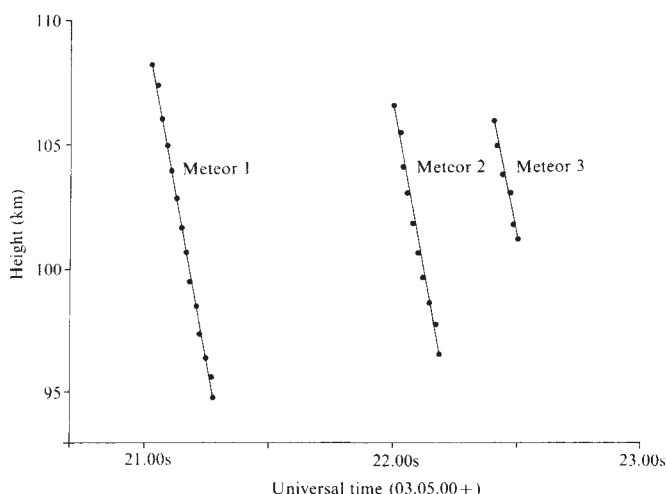


Fig. 1 The variation of meteor heights as a function of time.

with $M = 3 \times 10^{-4}$ kg, $T_c = 1,200$ yr (4×10^{10} s). The probability of collision with an interplanetary dust particle within 7×10^6 km of the Earth, for a meteor of mass 3×10^{-4} kg travelling at 58.5 km s $^{-1}$, is $\sim 3 \times 10^{-6}$. This is a small probability but still $\sim 10^5$ times higher than the chance probability of three meteors arriving so close together in time and direction.

Fechtig *et al.*¹⁰ have reported that the flux of dust particles in the mass range 10^{-12} – 10^{-18} kg observed from the HEOS 2 satellite within the Earth's magnetosphere, is ~ 100 times greater than the flux in interplanetary space. If a similar enhancement exists for masses of $\sim 10^{-10}$ kg, the collisional lifetime for a Perseid in this region would be reduced to $\sim 4 \times 10^8$ s. It takes $\sim 1,000$ s for a Perseid to cross the magnetosphere, thus the probability of a catastrophic collision with a magnetospheric dust particle is therefore $\sim 3 \times 10^{-6}$, which is comparable with the probability of a collision with a dust particle in interplanetary space.

Fechtig *et al.*¹⁰ also observed groups of particles from the HEOS 2 satellite which are believed to be high velocity ejecta from the impact of large meteoroids on the surface of the Moon. The flux of these particles was about one-tenth of the flux of dust particles in interplanetary space. During meteor showers the number of lunar ejecta increases. Alexander and Corbin¹¹ have calculated that the mass flux ejected from the lunar surface by the Perseid meteor shower is ~ 30 times that ejected by sporadic meteors. They have also shown that when the Moon is in the last quarter, a large percentage of the lunar ejecta are directed towards the Earth. At the time of our observation of 12 August 1977, the lunar phase angle from full Moon was 146° (in the last quarter). The flux of lunar ejecta near the Earth may, therefore, have been about three times the interplanetary dust flux. If these ejecta particles are concentrated in a region comparable in size to the Moon's orbit (4×10^5 km) then the probability of a catastrophic collision with a lunar ejecta particle is $\sim 5 \times 10^{-7}$.

The three Perseid meteors we observed ablating in the Earth's atmosphere, closely associated in space and time, are likely to be fragments of a larger Perseid meteoroid of mass 3×10^{-4} kg which broke up between $1,700$ and 7×10^6 km from the Earth's surface. The probability P_1 that they are fragments from a single meteoroid colliding with a dust particle in interplanetary space is $\sim 10^5$ times greater than the probability of a random association between the three meteors. The probability P_2 that the parent Perseid collided with a dust particle in the higher density dust environment of the Earth's magnetosphere is comparable with P_1 . The probability P_3 that the parent Perseid collided with a moon dust particle ejected when the Moon was bombarded with meteors from the Perseid shower, is $\sim$ one-sixth of P_1 .

If the Perseid broke up on collision with a dust particle we were lucky to observe this event. The chance of an individual meteor breaking up close to the Earth was of the order of a few in a million and we recorded less than 100 meteors during the 1977 Perseid meteor shower. During spectacular displays of meteors, when the parent comet passes close to the Earth, the meteor rate can rise by two or three orders of magnitude. The dust population in the magnetosphere and the density of lunar ejecta in near Earth space should rise and the chance of an individual meteoroid breaking up close to Earth on collision with a dust particle could increase up to a few in a thousand. In this case several events of the type we have described could be observed during the meteor display. The parent comet of the Perseid meteor stream (Swift–Tuttle, 1862 III) is expected to return to perihelion sometime in the next year or two. There has been speculation that the passage of the comet may give rise to a spectacular Perseid display. We may then be able to make further studies on the breakup of Perseid meteors in near Earth space.

We thank Dr O. Royrvik for observational assistance and Dr J. V. Jelley for use of his observing site at Harwell.

Received 28 August 1980; accepted 16 February 1981.

1. Cook, A. F. *NASA Contractor Rep.* No. 2109, 153–166 (1972).
2. Lovell, A. C. B. *Meteor Astronomy*, 131 (Oxford University Press, 1954).

3. Hughes, D. W. in *Cosmic Dust* (ed. McDonnell, J. A. M.) 151 (Wiley, Chichester, 1978).
4. Hughes, D. W. in *Cosmic Dust* (ed. McDonnell, J. A. M.) 164 (Wiley, Chichester, 1978).
5. Hawkes, R. L. & Jones, J. *Mon. Not. Roy. astr. Soc.* **185**, 727–734 (1978).
6. Paddack, S. J. & Rhee, J. W. *Geophys. Res. Lett.* **2**, 365–367 (1975).
7. Southworth, R. B. *Smithsonian Contr. Astrophys.* **7**, 299–303 (1963).
8. Moore, H. J. & Gault, D. E. *U.S. geol. Survey A. Progr. Rep.* 127–150 (1965).
9. Dohnanyi, J. S. *J. geophys. Res.* **75**, 3468–3493 (1970).
10. Fechtig, H., Grün, E. & Morfill, G. *Planet. Space Sci.* **27**, 511–531 (1979).
11. Alexander, W. M. & Corbin, J. D. *Submicron Lunar Ejecta in the Magnetosphere Associated with Meteor Showers* (COSPAR, Budapest, 1980).

Atomic motion on the surface of a cadmium telluride single crystal

R. Sinclair, T. Yamashita & F. A. Ponce

Department of Materials Science and Engineering,
Stanford University, Stanford, California 94305, USA

Cadmium telluride (CdTe) is an interesting semiconductor compound whose photoelectronic properties make it suitable for solar cell applications. We report here a study of defects in CdTe, carried out by lattice resolution imaging in a transmission electron microscope (TEM), during which observations were made of the motion of image spots at the edge of a specimen. As the spots are related to the projected atomic positions in the structure, this behaviour indicates atomic motion at the CdTe surface. Individual spot jumps occurred several times per second and one could watch the formation and decay of small crystalline facets. The possibility thus arises for surface studies, at the atomic level, in conjunction with direct crystal lattice imaging, allowing investigation of the interaction between surface atoms and their substrate.

CdTe has the zinc blende structure which is best imaged by high-resolution TEM with a $\langle 110 \rangle$ direction parallel to the electron beam. The projection of the structure in this orientation is shown in Fig. 1a. Adjacent columns of cadmium and tellurium atoms are separated by 1.6 Å, but their resolution is only achieved in very restrictive experimental circumstances^{1,2}. More commonly, bright spots are seen in the image (for example, Fig. 1b), the distribution of which corresponds to the arrangement of the projected lattice points. Such images are consistent with comparable images of elemental germanium and silicon which have the related diamond cubic structure^{3–5}. Theoretical image calculations, discussed later, establish the relationship between image spots and the crystal structure.

Microscopy was performed using a Philips EM400 TEM equipped with a high-magnification objective lens ($C_s = 1.1$ mm at 120 kV) and a LaB₆ electron source. This lens has no tilting capability. Therefore the single crystals were sectioned as closely as possible to the desired orientation. Thin foils were obtained by standard ion-milling procedures. A region of the specimen close to the perfect 110 zone axis orientation was found by translating the specimen in the microscope. The present images were taken by allowing 36 beams through the objective aperture (Fig. 1c) although the observations have been repeated with seven beam images. The magnification was $\times 800,000$, with typical exposure times of about 1 s.

In these experimental conditions, it was found that after an initial incubation period the single-crystal specimen deteriorated by the nucleation of new, small crystallites, apparently on the surfaces of the parent crystal. The deterioration time varied from one specimen to another and was typically several hours. Close examination of the specimen edge revealed that individual bright spots were moving about by a random jump process. Jumps occurred several times per second, usually by a distance equal to the atomic column separation. Not all edge spots were moving simultaneously; rather one or two would be moving back and forth along a crystal facet. Unfortunately the time both for transporting photographic

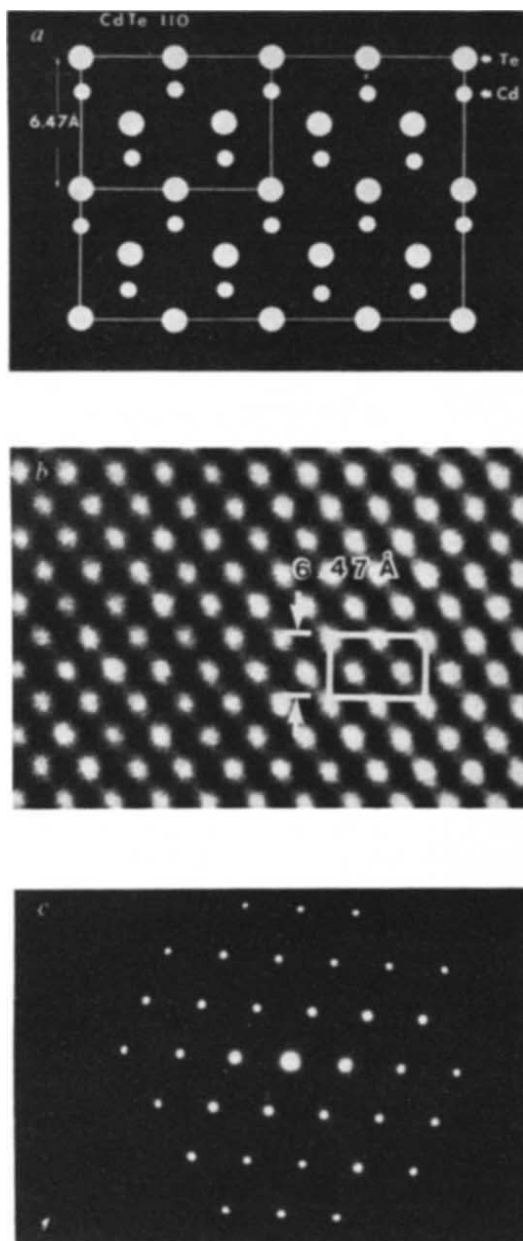


Fig. 1 *a*, Projection of the CdTe structure in the $\langle 110 \rangle$ orientation. *b*, Typical structure image of CdTe (atoms unresolved), with the unit cell outlined. *c*, Electron diffraction pattern showing the reflections used in image formation.

plates into position and for exposing the film was too long to allow the motion of an individual spot to be studied extensively. The jump rate could not be reduced to an acceptable level, for instance through decreasing the incident electron beam current density by defocusing the microscope condenser lens.

The motion of the surface spots is demonstrated here by the appearance of the specimen after intervals of several minutes. Thus it can be seen in Fig. 2 that the shape changes noticeably with time, although the crystal structure of the interior is always maintained. The general impression was that small facets were produced, the longer ones being parallel to $\{111\}$ -type planes. The bright spots at the edge move from one lattice site to another. When they become incorporated into the crystal, they extend it and the image is indistinguishable from the perfect crystal image. These observations, therefore, seem to indicate that atomic motion is occurring along the edge of the specimen, and probably across all surfaces. Those atoms which move on

the surface normal to the electron beam will coincide exactly with the underlying atomic columns if tetrahedral coordination is to be maintained. Therefore one would not detect directly motion of these latter atoms.

To interpret the image-spot motion the relationship between the image and the CdTe crystal structure must be understood. As the image comprises bright spots, the specimen must have finite thickness. Image calculations¹ indicate that the thickness is probably in the range 40–100 Å. Furthermore, to detect a bright image spot an atomic column of at least several unit cells is necessary. Thus the spots at the specimen edge do not arise from single atoms nor from single CdTe molecules but rather from a significant column of atoms probably of dimension similar to that of the adjacent bulk crystal. The movement of spots represents coordinated motion of a whole or partial atomic column but not the jumping of single atoms. For the observation to be made, the number of atoms that move must be sufficient to produce a detectable image spot. In the present case this is thought to be a column of about five atoms (~ 20 Å long).

The image spots may coincide with three positions in the projected structure depending on the particular combination of the experimental variables of microscope defocus, spherical aberration coefficient and specimen thickness¹. These positions are at the open channels in the structure, at the position of one species or at the midway point between the closest Cd–Te atomic columns. (Only the first and last occur in images of Ge and Si⁶.) However, the image spot motion is independent of its exact position. Thus the jump of the Cd–Te molecule 'A' in Fig. 3 would be seen as a spot moving whether the spot coincides with the molecule, one atom, or the adjacent column. The argument also holds for other possible migration mechanisms indicated in Fig. 3. Distinction of the various mechanisms would require quantitative measurements of the intensity of particular image spots, at known focus settings, including careful image recording by a television system.

The driving force for the present behaviour is presumably reduction in total surface area. Using a simple formulation for atomic jump frequency ($\nu_0 \exp -Q/kT$), it is found either that the activation energy (Q) for the present observations is 0.8 eV if they occur at room temperature, or that the specimen temperature is ~ 110 °C if the activation energy is assumed to be the sublimation energy of CdTe (1.03 eV)⁷. The latter is thought to represent the present situation. Thus, the observations are only made after the specimen temperature rises sufficiently due to heating by the incident electron beam. As exposure continues, the motion becomes more rapid with more drastic specimen deterioration. Variation between specimens arises from their exact geometry and from the degree of contact between the specimens and a supporting metal grid. Complete isolation of the CdTe is found to result in extremely rapid specimen destruction. The slow response of the jump rate on reducing the beam current density indicates that the atomic motion is not assisted by direct electron–atom collisions (the electron energy is well below the threshold for Cd or Te displacement) or by ionization effects, and probably arises from a sluggish heat dissipation from the active area. A more sophisticated analysis would need to take into account the mechanism of coordinated atomic column motion.

Whether these results are confined to CdTe or represent a more general phenomenon, previously undetected, remains to be seen. Certainly the absence of a specimen contamination film in the present work is a factor contributing to the observations. CdTe has the lowest sublimation energy of the common semiconducting materials⁷. If the present ideas are correct then equivalent observations may be made on other similar solids if higher temperatures are achieved. For example, jump frequencies several times per second would be expected at 600 °C for silicon ($Q = 2.32$ eV), 340 °C for gallium arsenide ($Q = 1.63$ eV) and 260 °C for cadmium sulphide ($Q = 1.42$ eV)⁷. *In situ* specimen heating, together with lattice resolution, would be necessary to detect these phenomena which should soon be possible on commercial TEMs.

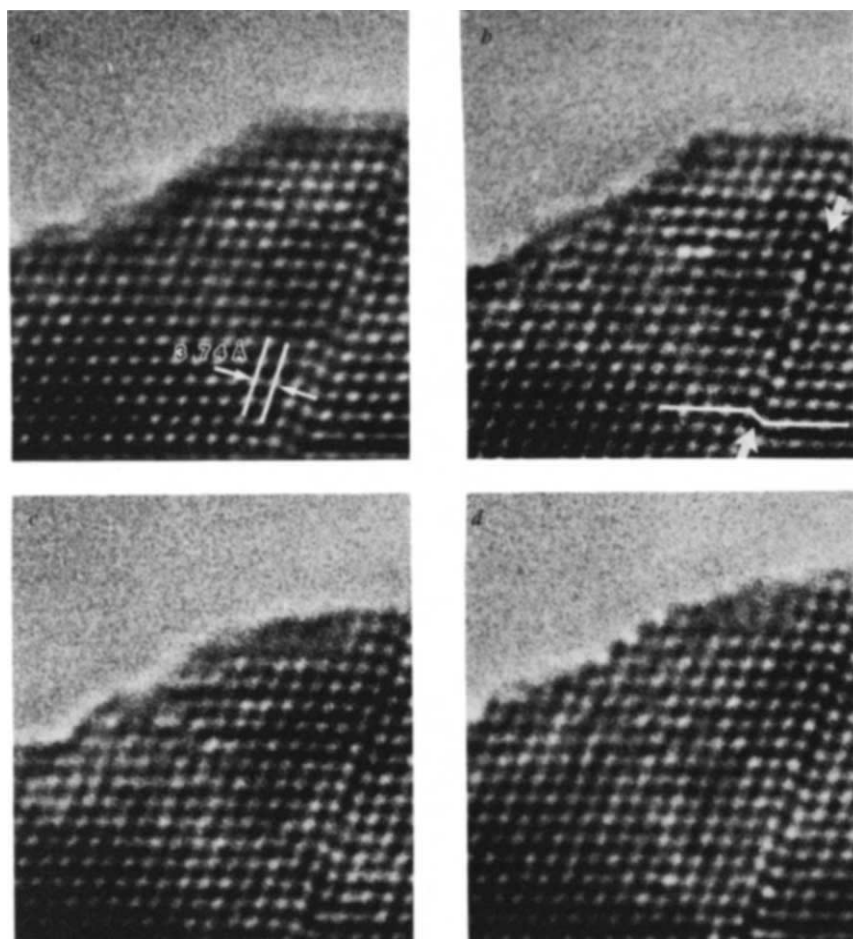


Fig. 2 Series of structure images of the same region near the specimen edge taken several minutes apart. The stacking fault (arrowed in *b*) can be used as a reference mark for each photograph. The specimen shape and its spot distribution change noticeably with time.

Direct investigations of atomic motion by electron microscopy have previously only been made without imaging the substrate. Scanning transmission microscopy results of heavy-metal atoms on amorphous carbon^{8,9} are well known, with complementary reports using fixed-beam TEM by dark field¹⁰ or bright-field imaging of atoms on graphite¹¹. Field-ion microscopy allows the imaging of both surface and substrate^{12,13} albeit of limited areas on a restricted number of materials. The present observations are thought to be the first using fixed-beam TEM, combining direct imaging of the moving species with the crystal structure of the parent material. However, unlike these other studies, they represent motion of a small group of atoms, by an undetermined mechanism, rather than of single atomic jumps.

The observations reported here have been interpreted as representing the migration of atomic columns (comprising at least five atoms and probably more) due to thermal activation induced by beam heating. A more detailed description of the

mechanism of this behaviour requires more rapid recording (for example, at television rate), extensive image calculations relating spot brightness to column thickness and a more extensive theoretical model of coordinated atomic jumps. Although this work is preliminary it has important implications. Thus with careful experimentation it should be possible to study directly self-migration on CdTe and determine the associated atomic mechanisms. If, as we suspect, the observations are not confined to CdTe, then atomic motion studies on this important class of materials are feasible. Finally, other surface-related phenomena may also be investigated, for example, the action of foreign species (such as metal atoms) and the influence of crystallography and crystal defects on the atomic behaviour. Such experiments would provide important information for understanding how the surfaces of such solids may influence their properties and their technological applications.

Financial support from the Alfred P. Sloan Foundation (R.S.), Xerox Corporation (T.Y.), Basic Energy Sciences Division of DOE (F.A.P.) and the NSF (equipment grant) is gratefully acknowledged.

Received 27 August 1980; accepted 27 January 1981.

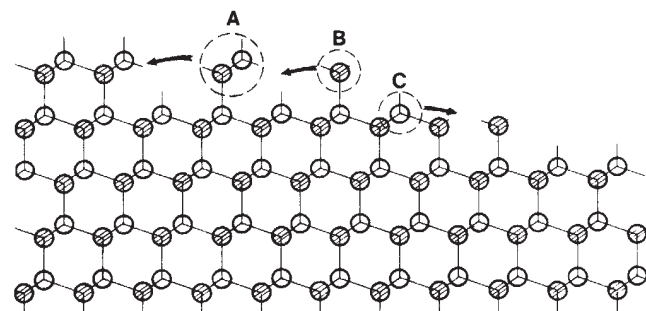


Fig. 3 Schematic model of possible mechanisms for surface atomic migration in CdTe.

1. Yamashita, T., Ponce, F. A., Pirouz, P. & Sinclair, R. (in preparation).
2. Yamashita, T. 38th *a. Proc. EMSA*, 168-169 (1980).
3. Spence, J. C. H., O'Keefe, M. A. & Kolar, H. *Optik* **49**, 307-323 (1977).
4. Rez, P. & Krivanek, O. L. *Proc. 9th Int. Congr. on Electron Microscopy* **1**, 288-289 (1978).
5. Desseaux, J., Renault, A. & Bourret, A. *Phil. Mag.* **35**, 357-372 (1977).
6. Olsen, A. & Spence, J. C. H. *Phil. Mag.* (in the press).
7. Harrison, W. A. *Electronic Structure and the Properties of Solids*, 176 (Freeman, San Francisco, 1980).
8. Isaacson, M., Ohtsuki, M. & Utlaut, M. *Introduction to Analytical Electron Microscopy* (eds Hren, J. J., Goldstein, J. I. & Joy, D. C.) 343-368 (Plenum, New York, 1979).
9. Crewe, A. V. *Chem. Scr.* **14**, 17-20 (1978-79).
10. Hashimoto, H., Yokota, Y., Takai, Y., Endoh, H. & Kumao, A. *Chem. Scr.* **14**, 125-128 (1978-79).
11. Iijima, S. *Optik* **48**, 193-214 (1977).
12. Ehrlich, G. & Hudda, F. G. *J. chem. Phys.* **44**, 1039-1049 (1966).
13. Tsong, T. T. *Chem. Scr.* **14**, 7-15 (1978-79).

High-voltage, high-resolution lattice images of dolomite

D. J. Barber

Physics Department, University of Essex, Colchester CO4 3SQ, UK

L. A. Freeman & David J. Smith

High Resolution Electron Microscope, University of Cambridge, Cambridge CB2 2QH, UK

Dolomite, $\text{CaMg}(\text{CO}_3)_2$, which belongs to the rhombohedral group of carbonates, is an important mineral commonly found in sedimentary and metamorphic environments. Ideally, it has an ordered trigonal structure ($R\bar{3}$) with planes containing atoms of only Ca, or only Mg, alternating and being interleaved with the planes containing the carbonate groups. All these planes are perpendicular to the 3-fold axis. As the microstructural deformation observed in many natural dolomites of low and medium metamorphic grade could provide an insight into the geological history, attention has recently been directed to electron microscope studies¹⁻⁴ of the mineral. Unfortunately, at a typical accelerating voltage of 100 kV, dolomite is highly susceptible to damage by the electron beam, whereas, at a voltage of 1 MV, little electron damage is noticeable¹. We report here observations of dolomite with the Cambridge University 600-kV high-resolution electron microscope^{5,6}. Although electron beam damage is still observed at the intermediate operating voltages of this microscope, it has been possible, with care, to resolve directly details of crystal lattices in two dimensions. Thus, direct confirmation has been obtained of postulated models for the common types of stacking faults observed in dolomite.

Much of the recent interest in dolomite stems from three observations:

(1) Single crystals of stoichiometric dolomite exhibit an unusual increase in mechanical strength with temperature up to $\sim 800^\circ\text{C}$ (refs 7, 2, 4), which seems to be a consequence of the detailed atomic displacements accompanying dislocation motion through the ordered structure.

(2) Attempts to synthesize dolomite in conditions similar to those in which sedimentary dolomites seem to have formed have failed^{8,9}. Moreover, although the ordered phase is more stable than disordered dolomite, the latter can exist for geologically long times. 'Geologically young' dolomites are mostly imperfectly ordered and calcium rich; embryonic analogues of the abundant, more perfect dolomites found in old environments have not been discovered.

(3) Recent work has shown that local disorder can be produced in near-perfect single crystal dolomite by slip on the (0001)^{2,4} and $\{01\bar{1}2\}$ planes⁴. X-ray structure refinements on Eugui dolomite which was thermally disordered and then quenched in a laboratory experiment indicate¹⁰, however, that other forms of disorder also occur in which some of the carbonate groups (which should ideally be oriented at 180° about the 3-fold axis with respect to the groups in adjacent carbonate planes) are misoriented by 180° . In addition, electron microscopy of non-stoichiometric sedimentary dolomites¹¹ has shown the existence of a new type of superstructure which effectively doubles the length of the a axis.

All these observations indicate the value of gaining direct information about types of disorder in dolomite. An obvious approach would involve lattice imaging techniques with a high-resolution electron microscope although, as mentioned above, dolomite is considerably electron sensitive because of local heating and ionization damage caused by the electron beam. Indeed, previous efforts at 100 kV have barely resolved the 0.54-nm (0003) lattice fringes in one-dimension only¹⁰. However, results have been improved by raising the beam

energy, which effects an appreciable decrease in the sensitivity of dolomite to beam damage, and by applying a variation of the so-called 'minimal exposure technique'¹² whereby tilting of the dolomite crystal to the required orientation is done using a focused beam of very low intensity on a small area of crystal. After focusing and correction of astigmatism, images are recorded of crystal defects and other interesting morphological features in nearby areas, previously unexposed to the intense electron beam. Local bending of the crystal lattice sometimes occurred but accurate alignment of the crystal normal and the zone axis projection was generally retained so that two-dimensional crystal structure images could be obtained.

The two-dimensional lattice image in Fig. 1 shows stacking faults produced by slip on the $\{10\bar{1}2\}$ planes in a single crystal of a stoichiometric and fairly pure dolomite (Eugui) which was experimentally deformed at 20°C under a confining pressure of 0.7 GPa. The electron beam is exactly parallel to a $\langle 10\bar{1}1 \rangle$ zone axis, as indicated by the selected area diffraction pattern. Two sets of $\{1\bar{1}01\}$ and one set of $\{10\bar{1}2\}$ lattice planes are resolved, with spacings of 0.40 and 0.37 nm respectively.

Figure 2a shows the structure image of a thin crystal of the same Eugui dolomite which was experimentally deformed at 420°C under confining pressure in an orientation which promoted slip on the (0001) planes. In this case the electron beam is exactly parallel to a $\langle 11\bar{2}0 \rangle$ direction. $\{1\bar{1}01\}$, $\{10\bar{1}2\}$, $\{1\bar{1}04\}$ and (0003) planes are parallel to the beam, as can be seen from Fig. 2b, which shows the atomic structure when viewed in this projection. The orientations of the carbonate groups, whose planes are strictly normal to the projection plane, are given by black triangles in this representation. A basal stacking fault, which results from slip on the (0001) $\langle 1\bar{2}10 \rangle$ macroscopic system, appears in Fig. 2a. The fault is apparently bounded by $1/3[1\bar{1}00]$ partial dislocations, as previously deduced from diffraction contrast experiments¹³, but it is broader in the $\langle 0001 \rangle$

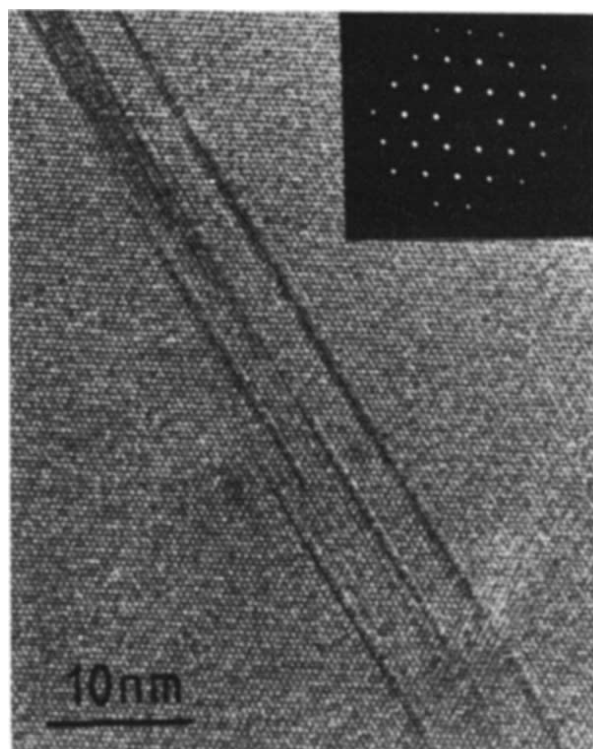


Fig. 1 Lattice image from a moderately thick crystal of dolomite in a $\langle 10\bar{1}1 \rangle$ projection showing a series of deformation-induced planar stacking faults on the $\{10\bar{1}2\}$ planes, together with the corresponding electron diffraction pattern (inset). The lattice planes imaged are $\langle 10\bar{1}1 \rangle$, $\langle 10\bar{1}2 \rangle$ and $\langle 01\bar{1}1 \rangle$; the faults do not produce apparent offsets of these planes.

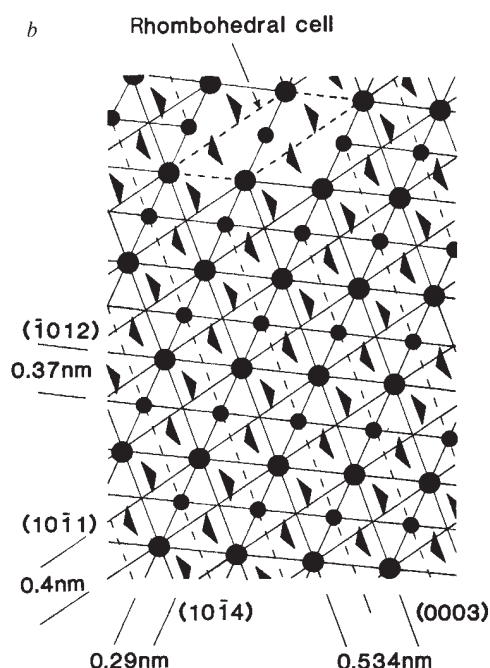
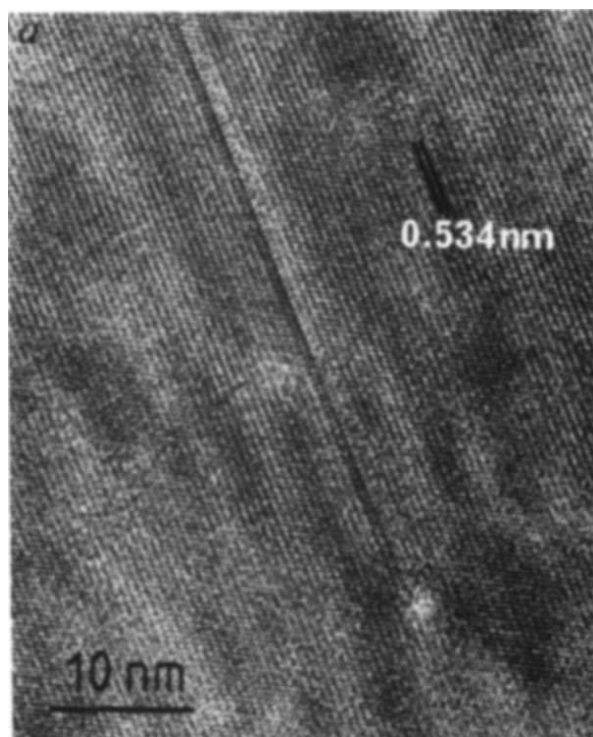


Fig. 2 *a*, Structure image of a thin crystal of dolomite taken in a $\{11\bar{2}0\}$ projection; the 0.534-nm (0003) basal plane spacing is indicated. A deformation-induced basal stacking fault is visible. Confirmation of postulated models for such stacking faults are directly obtainable. *b*, Representation of a section through a $\{11\bar{2}0\}$ plane of dolomite, showing alternating layers of Ca and Mg atoms, separated by planes of CO_3 groups parallel to the basal planes (0001). The plane of each CO_3 group is actually exactly parallel to (0001) but small (tilted) triangles are used to denote the 180° rotation between groups in adjacent planes.

direction than might be expected from elementary considerations. A more complete analysis and interpretation of (0001) and $\{1012\}$ stacking faults will be presented elsewhere. The disparity in cation sizes is apparent in the structure image of Fig. 2*a*, although, as might be expected, the carbonate groups are not seen in detail.

Clearly, high-voltage, high-resolution electron microscopy is a technique well suited to analysis of order/disorder problems in dolomite. Furthermore, use of this method is expected to add greatly to our knowledge of defect structures in non-stoichiometric and disordered specimens of other carbonates. Sequential images of dolomite which has undergone prolonged electron irradiation have also been obtained, from which it is believed that the nature of the electron damage process will be established. Our preliminary results suggest that the damage propagates along preferred crystallographic directions in common with earlier observations of phthalocyanines¹⁴. This damage ultimately leads to the decomposition of the carbonate to the oxides, as with calcite^{15,16}.

The Cambridge University 600-kV high-resolution electron microscope has been built as a joint project between the Cavendish Laboratory and the Department of Engineering with major financial support from the SRC. We thank the latter for continuing support, R. A. Camps for assistance with preparation of the microscope and H. C. Heard, R. J. Reeder and H. R. Wenk for provision of samples and valuable discussions.

Received 16 December 1980; accepted 18 February 1981.

1. Barber, D. J. *Tectonophysics* **39**, 193 (1977).
2. Heard, H. C., Wenk, H. R. & Barber, D. J. *EOS* **59**, 249 (1978).
3. White, J. C. & White, S. H. *Tectonophysics* **66**, 35 (1980).
4. Barber, D. J., Heard, H. C. & Wenk, H. R. *Phys. Chem. Miner.* (in the press).
5. Nixon, W. C. *et al.* in *Developments in Electron Microscopy and Analysis 1977* (ed. D. L. Misell) 13–16 (Institute of Physics, Bristol & London, 1977).
6. Cosslett, V. E. *Proc. R. Soc. A* **370**, 1 (1980).
7. Higgs, D. V. & Handin, J. W. *Bull. geol. Soc. Am.* **70**, 245 (1959).
8. Garrels, R. M., Thompson, M. E. & Siever, R. *Am. J. Sci.* **258**, 402 (1960).
9. Berner, R. A. in *Principles of Chemical Sedimentology* (McGraw-Hill, New York, 1971).
10. Reeder, R. J. thesis, Univ. California, Berkeley (1979).
11. Reeder, R. J. & Wenk, H. R. *Geophys. Res. Lett.* **6**, 77 (1979).
12. Williams, R. C. & Fisher, H. W. *J. molec. Biol.* **52**, 121 (1970).
13. Barber, D. J., Heard, H. C., Paterson, M. S. & Wenk, H. R. *Nature* **269**, 789 (1978).
14. Murata, Y., Fryer, J. R., Baird, T. & Murata, H. *Acta crystallogr.* **A33**, 198 (1977).
15. Burrage, B. J. & Pitkethly, D. R. *Phys. Status Solidi* **32**, 399 (1969).
16. Towe, K. M. *Nature* **274**, 239 (1978).

Near-surface turbulence measurements in a lake

T. M. Dillon, J. G. Richman, C. G. Hansen & M. D. Pearson

School of Oceanography, Oregon State University, Corvallis, Oregon 97331, USA

One of the most important but least understood regions of oceans and lakes is the near-surface zone. Exchanges of energy and momentum between the atmosphere and the waters below are controlled by turbulent mixing processes in the surface zone. Major advances in understanding these processes have been hampered by a lack of observations of turbulence near the surface. We have recently constructed a probe which can be used to estimate an important variable in theories of turbulence, the rate at which the kinetic energy of turbulence is dissipated, and have found a surprising though simple result: to a first approximation, scaling laws devised to describe dissipation in the atmospheric boundary layer may also be applicable in the water near the surface.

The probe is a miniaturized version of a high-frequency temperature profiler used previously in upper ocean studies¹. The probe body is a 2.5-cm diameter cylinder, 70-cm long, containing thermistor amplifiers and a pressure transducer. A toroidal buoyancy element 10 cm in diameter and 25-cm long jackets the upper probe body, and a vented drag plate 30-cm in diameter is attached to the rear. The probe is slightly buoyant in water and has a large static stability. A thermistor with a frequency response correctable to 30 Hz is mounted on the nose of the probe. Ballast is attached with a pressure-sensitive release mechanism so that after launching, the probe sinks to a pre-determined depth, releases ballast and rises to the surface at

10 cm s⁻¹. The probe is launched from a small boat, and typically rises 15–20 m away from the launch point, so that water sampled by the probe is unaffected by the boat.

Estimating the dissipation rate from high-frequency temperature measurements depends on finding the wavenumber (inverse length scale) of the smallest appreciable temperature fluctuations^{2–4}. Briefly, the method assumes a competition between the tendency of turbulent stirring to increase temperature gradients and the tendency of molecular diffusion to smooth gradients. The wavenumber at which diffusion can overcome stirring and homogenize the fluid, the Batchelor wavenumber k_B , is a function of the kinetic energy dissipation rate per unit mass, ϵ : $k_B = (\epsilon \nu^{-1} D^{-2})^{1/4}$, where ν is the kinematic viscosity and D the molecular diffusivity for temperature. Wavenumber spectra are calculated from temperature-gradient fluctuations, and by least-squares fitting to the Batchelor spectrum^{2,5,6}, k_B and hence ϵ may be found.

The validity of the method has been tested in the nepheloid layer near the ocean floor. Measurement of the stress on the ocean floor due to near-bottom currents⁷ enables the dependence of the kinetic energy dissipation rate on height above the floor to be predicted. Simultaneous measurement of the temperature microstructure was made and the Batchelor scale method was used to calculate the dissipation rate above the ocean bottom^{8,9}. Comparison of the measurement with the prediction yields excellent agreement⁴.

Measurements were made in Green Peter Reservoir, a small impoundment east of Corvallis, Oregon. The launch site had an unobstructed fetch of 3 km over water, and the depth at anchorage was 50 m. During the experiment, winds averaged 4.8 m s⁻¹. Surface waves averaged 10–20 cm in height, and occasional white-capping was observed. The temperature profile (Fig. 1) displayed an active surface mixing layer to 6 m in depth. A strong seasonal thermocline was at 7 m.

Nine casts were made through the mixing layer, and temperature gradient spectra were computed over 50-cm segments of water. From the surface to 1-m depth, the turbulence was so intense that the frequency response of the thermistor was insufficient to resolve the smallest length scales on half of the segments, and so the dissipation rate could not be reliably resolved; in these cases, we underestimate the dissipation by

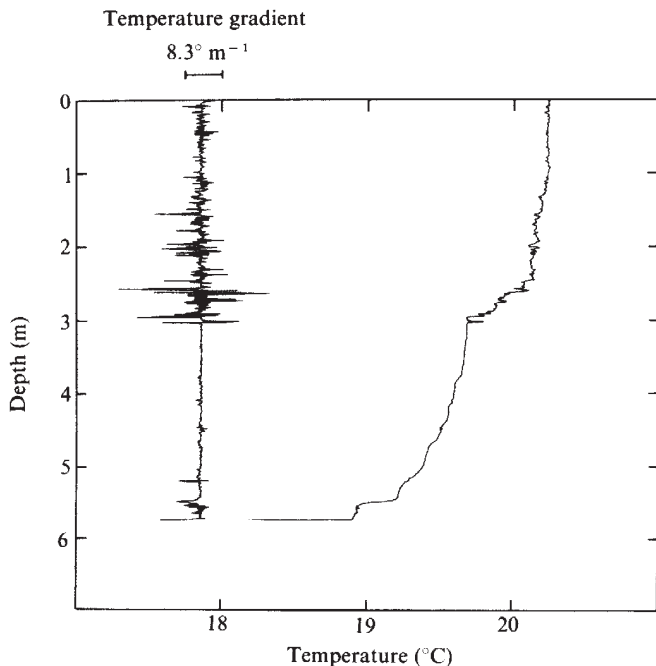


Fig. 1 Temperature (right) and temperature gradient (left) profiles for a single cast in the near-surface layer at Green Peter Reservoir, 24 July 1980. Thermocline is at 6-m depth.

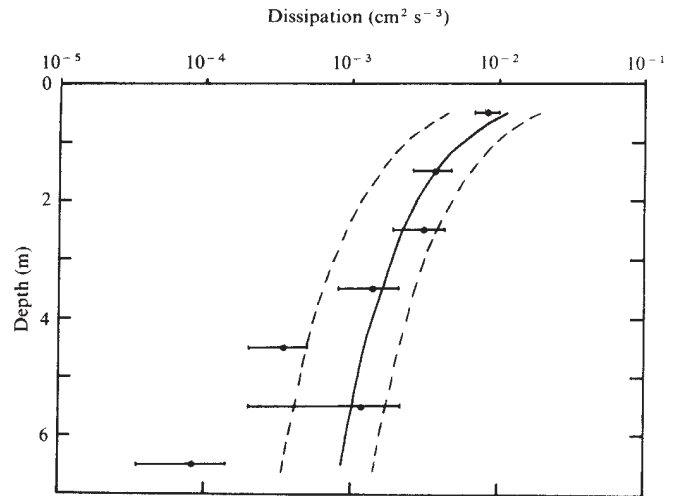


Fig. 2 Kinetic energy dissipation rate per unit mass in the near-surface layer (●). Error bars are the standard error of the mean. Solid line is $(\tau/\rho)^{3/2} \kappa^{-1} Z^{-1}$ calculated from mean winds, dotted lines are $(\tau/\rho)^{3/2} \kappa^{-1} Z^{-1}$ calculated from highest and lowest winds. Dissipation falls dramatically in the thermocline at 6 m.

setting it equal to the largest resolvable value. Below the first metre, the dissipation rate decreases, and was resolvable on 90% of the segments. At the base of the layer, the dissipation rapidly falls by an order of magnitude. Averages of the dissipation over the nine casts were calculated in 1-m intervals (Fig. 2).

No detailed theoretical framework for understanding turbulence in a wind driven surface mixed layer has yet received wide acceptance. Attempts at detailed modelling have been made¹⁰ but the complexities of possible interactions of wind, waves, breaking waves and spray, as well as the lack of near-surface observations, have kept most theories of the aquatic mixed layer at a primitive level. In contrast, the theory of the atmospheric boundary layer has been well developed¹¹. It is tempting to ignore the complex nature of the air-sea interaction and apply to the aquatic mixed layer the same framework which holds on the atmospheric side of the boundary. The atmospheric boundary layer prediction in an unstratified flow for variation of dissipation rate with distance Z from the boundary¹¹ is $\epsilon = (\tau/\rho)^{3/2} \kappa^{-1} Z^{-1}$, where κ is the von Karman constant and ρ the water density. Calculating the surface stress during our measurements by the drag law $\tau = \rho_a C_D \bar{U}^2$ (ρ_a is the air density, C_D (ref. 12) is the drag coefficient with value 1.3×10^{-3} , and $\bar{U}$ is the mean wind speed at a height of 5 m), we find remarkably close agreement with the atmospheric boundary-layer prediction. A least-squares fit to the form $\epsilon = AZ^n$ yields $n = -1.03 \pm 0.05$. Within experimental error, the dissipation is inversely proportional to the depth. Testing the magnitude of the dissipation, we find the average value $A = \epsilon Z = 0.51 \pm 0.09 \text{ cm}^3 \text{ s}^{-3}$ which is within experimental error of the independently measured value $(\tau/\rho)^{3/2} \kappa^{-1} = 0.53 \text{ cm}^3 \text{ s}^{-3}$. Exhaustive measurements of ϵ will be required to determine whether or not the atmospheric boundary layer prediction holds to more than a first approximation.

It has been suggested^{13–15} that the near-surface layer may exhibit some of the properties of a constant-stress layer, but the experimental evidence has been too fragmentary to draw firm conclusions. It is premature to suggest that the scaling will hold at all times. At large wind speeds, vigorously breaking surface waves may be an important source of turbulence, and have scaling laws different from the case of small waves and low winds.

This work was supported by the NSF under contracts OCE 7825667 and OCE 8018444, the Office of Naval Research under contract ONR 00014-79-C-0004, and the Oregon State University School of Oceanography Summer Work Program.

Received 2 September 1980; accepted 23 January 1981.

1. Caldwell, D. R., Wilcox, S. D. & Matsler, M. *Limnol. Oceanogr.* **20**, 1035–1047 (1975).
2. Dillon, T. M. & Caldwell, D. R. *J. geophys. Res.* **85**, 1910–1916 (1980).
3. Caldwell, D. R., Dillon, T. M., Brubaker, J. M., Newberger, P. A. & Paulson, C. A. *J. geophys. Res.* **85**, 1917–1924 (1980).
4. Caldwell, D. R., Chriss, T. M., Newberger, P. A. & Dillon, T. M. *J. geophys. Res.* (in the press).
5. Batchelor, G. K. *J. Fluid Mech.* **5**, 113–133 (1959).
6. Gibson, C. H. & Schwarz, W. H. *J. Fluid Mech.* **16**, 365–384 (1963).
7. Caldwell, D. R. & Chriss, T. M. *Science* **205**, 1131–1132 (1979).
8. Newberger, P. A. & Caldwell, D. R. *J. geophys. Res.* (in the press).
9. Newberger, P. A. & Caldwell, D. R. *J. geophys. Res.* (in the press).
10. Kundu, P. K. *J. phys. Oceanogr.* **10**, 220–236 (1980).
11. Lumley, J. L. & Panofsky, H. A. *The Structure of Atmospheric Turbulence* (Wiley, New York, 1964).
12. Amoroso, J. & Devries, J. J. *J. geophys. Res.* **85**, 433–442 (1980).
13. Jones, I. S. F. & Kenney, B. C. *J. geophys. Res.* **82**, 1392–1396 (1977).
14. Bye, J. A. T. *Limnol. Oceanogr.* **10**, 451–458 (1965).
15. Phillips, O. M. *The Dynamics of the Upper Ocean* 2nd edn (Cambridge University Press, 1977).

Geomagnetic induction on a transatlantic communications cable

L. V. Medford, A. Meloni,* L. J. Lanzerotti & G. P. Gregori†

Bell Laboratories, Murray Hill, New Jersey 07974, USA

The phenomenon of geomagnetically-induced currents in long conductors has been recognized since the mid-nineteenth century, when the use of the telegraph became prevalent in North America and Europe¹. Powering systems of telegraph and telephone cables are now engineered to eliminate large current surges, although the occasional significant geomagnetic disturbance, such as those produced by the February 1958², August 1972³ flare events, can still be a problem. However, induced voltages can be used, with complementary geophysical and geological data, to deduce features of the Earth's crust and upper mantle^{13,14}. Improved scientific understanding of magnetospheric phenomena means that the geomagnetic induction process in long conductors, and thus in the Earth are better understood. For example, the cable outage problem associated with the August 1972 geomagnetic storm interval was qualitatively attributed not to the 'line' currents associated with normal auroral current systems but rather to the significant enhanced currents flowing on the magnetosphere boundary at the time of an extreme compression of this boundary in greatly enhanced solar wind conditions³. We have measured the geomagnetic induction phenomenon on modern-day long cables to get better assessments of the induced currents in various types of geomagnetic variations. Our initial work was concerned with induction in the transatlantic telecommunications cable no. 6 (TAT-6) which is laid between Green Hill, Rhode Island (41.4° N, 71.7° W geographic) and St Hilaire de Riez, France (46.7° N, 1.9° W)⁴. We concentrate quantitatively here on induction by the solar-induced quiet day currents (S_q) in the ionosphere and find that calculated values agree well with observed values of S_q .

The solid line in Fig. 1 shows the approximate route of the TAT-6 cable across the North Atlantic (length 6,300 km). This cable is powered at both Green Hill and St Hilaire by constant current power supplies with opposite polarity at each end producing a current flow from west to east through the cable and repeaters^{5,6}. These repeaters are spaced approximately every 10 km along the route of the cable⁷. The cable is biased at a nominal 5,200 V positive with respect to local ground in North America and 5,200 V negative with respect to local ground at St Hilaire with a maximum capability of 7,500 V from each supply. The present measurements from the Green Hill terminal monitored the geomagnetically-induced currents in the powered

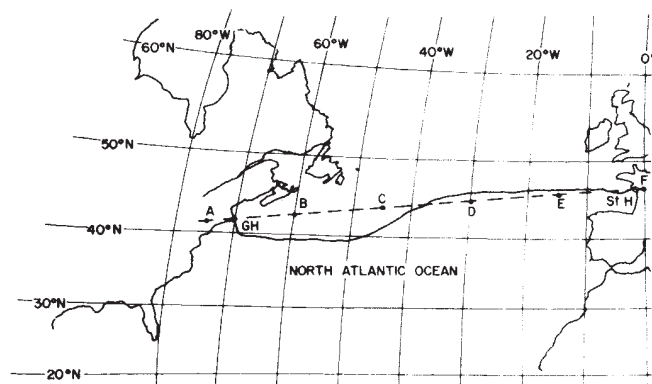


Fig. 1 Approximate location of TAT-6 cable across the North Atlantic (solid line). Model cable path (dashed line).

centre conductor by measuring changes in the conductor to local ground voltage. The voltage drop to ground at St Hilaire was not monitored in these first tests, but will be investigated in future work. To a first order, induction in the cable may be expected to produce equal and opposite voltage changes at the two ends of the system.

The constant current flow (nominal 657 mA) along the centre conductor can be pictured as coming from two sources at each cable end: one being the active power supply and the other the induced voltage produced by the rate of change with time of the geomagnetic field across the cable circuit. Thus, a measure of the time dependence of the externally-induced currents can be obtained by measuring the changes in the output voltage of the constant current power supply. The total d.c. resistance of the centre conductor and repeaters (under power) is $\sim 16 \text{ K}\Omega$. The time constant of the conductor produced by the distributed cable capacitance and the resistance (end to end) is $\sim 10 \text{ s}$. The dashed line in Fig. 1 shows the model cable path that was taken for the modelling of the S_q field in the calculations discussed below.

In addition to monitoring the cable power supply voltage, the Earth's magnetic field in three geomagnetic components (H (north-south), D (east-west), Z (vertical)) was measured at Green Hill between 1 and 18 May 1980. The cable voltage data and the magnetometer data were individually recorded on magnetic tape using a microprocessor-based low-power digital data acquisition system with 15-bit resolution. The three components of the magnetic field and the voltage data were digitized at 2-s intervals. Timing (absolute and relative) was

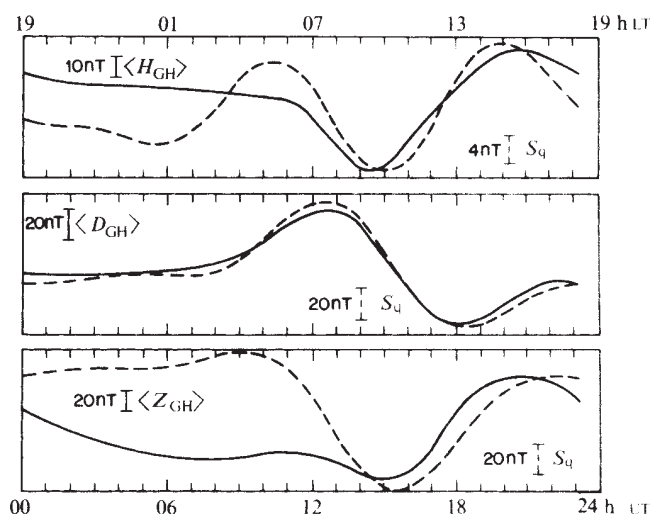


Fig. 2 Average of five quiet day magnetic field variations (solid line). Calculated magnetic field variations produced by S_q ionospheric current systems (dashed line).

Permanent addresses: * Istituto Nazionale di Geofisica, Osservatorio Geofisico Monte Porzio, 00040 Rome, Italy; † Istituto di Fisica dell' Atmosfera, CNR, 00144 Rome, Italy.

accurate in each data system to better than 1 s during the period of the measurements.

The time during which the cable measurements were made was relatively quiet geomagnetically⁸. Five of the most quiet days were selected to represent the S_q variation measured at Green Hill. These days, and their daily sum global geomagnetic activity indices ΣK_p , were: 3 May, -6; 4 May, -9; 5 May, -14; 17 May, +5; 18 May, -8. The averages of these five quiet days in the three magnetic field components measured at Green Hill are shown as solid lines in Fig. 2. These averages are obtained from the hourly mean values. The time dependences during a day show typical S_q patterns for this geomagnetic latitude, except for the local morning intervals in the H - and Z -components.

The dashed lines in Fig. 2 show the computed S_q curves expected on the basis of the spherical harmonic analysis of the potential for the daily variation by Chapman and Bartels⁹. We used the harmonic coefficients of Matsushita¹⁰ for location B on the model cable path in Fig. 1. The fit of the calculated S_q patterns using location B is quite good in the east-west component but is not good for the local morning observations in either the north-south or the vertical components. In addition, the magnitude of the S_q variation observed in the local afternoon in the north-south component is about a factor of two larger than that predicted by the model S_q calculations. These disagreements in the north-south and vertical components are

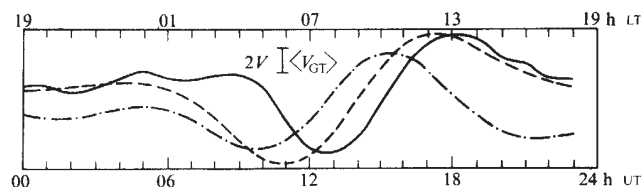


Fig. 3 Average of cable voltage variation for the five geomagnetic quiet days (solid line). Calculated induced voltage from all three magnetic components produced by S_q current system (dashed line). Calculated induced voltage from the horizontal magnetic component produced by S_q current system (dot-dashed line).

partially due to the closeness of the measurement site (~ 3 km) to the North Atlantic Ocean as well as to uncertainties in the spherical harmonic coefficients used in the S_q calculations. Note that data (not shown here) from our magnetometer station at Durham, New Hampshire, (44° N, 72° W geographic), located ~ 30 – 35 km from the ocean, show better agreement in the north-south component for the local morning S_q variations than do the Green Hill data.

The solid line in Fig. 3 shows the measured average variation of the cable voltage as a function of local time for the same five days as the magnetic field data of Fig. 2. This average was made using the hourly mean values and computing from them the five day average, in the same manner as was done for the S_q curves of Fig. 2. The relative expected induced voltage from induction by the S_q current system was modelled by considering the universal time (UT) change in the instantaneous S_q magnetic field pattern observed across the cable. To simplify the analysis, the S_q patterns at six locations (A–F in Fig. 1) were used for each hour of UT; the time derivative of the S_q magnetic field pattern was calculated at each of the six points. These six voltages ($V = -d\Phi(B)/dt$, where $\Phi = \mathbf{B} \cdot \mathbf{nA}$ is the magnetic flux through the surface S with area A) were averaged to represent the induced voltage across the cable at a given time. This induced voltage is of opposite polarity to the measured voltage. These calculated relative voltages (with changed signs) are superimposed on the measured voltage in Fig. 3. The dashed line is the calculated relative induction considering the total magnetic field changes from the S_q magnetic field pattern in all three components; that is, the total field contribution. The dot-dashed line is the calculated relative voltage considering only the horizontal component. The relative behaviour of this former calculated profile is a reasonably good approximation (with a 1- to 2-h

phase shift) to the observed voltage. (Induction currents from possible tidal^{2,11,13,14} and ocean current¹² effects are not considered important for this work^{2,11}.) A better agreement in terms of phase is obtained by assuming that the induction process that changes the Green Hill voltage occurs closer to the Green Hill end, rather than along the entire length of cable. Calculations of the S_q voltage pattern from points A–D (Fig. 1) show no phase shift. Physical understanding of this effect could be obtained by simultaneous measurements at both Green Hill and St Hilaire.

It is not possible to compare the magnitudes of the calculated voltage values with the observed S_q induced voltage without making some assumption as to the area of the Earth through which the induction process operates. An approximate equality between the relative calculated amplitudes and the observed induced voltage (~ 5 V) from the S_q pattern would imply an area of $\sim 10^{12}$ m² influenced by the induction.

Future work on induction by shorter-period geomagnetic variations should result in a better understanding of the geomagnetic induction phenomenon in varying geomagnetic conditions as well as at the different Earth locations over which a cable is laid.

We thank S. T. Brewer, B. H. Hamilton, and G. R. Ostby for advice and technical assistance, also G. Hailey and R. Foran for their support and help with our instrumentation at the Green Hill Terminal, and H. Wilhelmson for his careful reading of the manuscript.

Received 29 October 1980; accepted 28 January 1981.

1. Barlow, W. H. *Phil. Trans. R. Soc.* 61 (1849).
2. Axe, G. E. *Post Off. Electr. Engrn. J.* 61, 37 (1968).
3. Anderson, C. W., III, Lanzerotti, L. J. & MacLennan, C. G. *Bell Sys. Tech. J.* 53, 1817 (1974).
4. Ehrbar, R. D., Ford, A. E. & Gerbier, G. *Bell Sys. Tech. J.* 57, 2313 (1978).
5. Calkin, E. T., Golito, I., Schatz, W. J., Schroeder, R. E. & Shull, D. S. *Bell Sys. Tech. J.* 57, 2497 (1978).
6. Hamilton, B. H. *IEEE Trans. Ind. Appl.* IA-12, 378 (1976).
7. Brewer, S. T., Easton, R. L., Soulier, H. & Taylor, S. A. *Bell Sys. Tech. J.* 57, 2319 (1978).
8. *Solar Geophysical Data*, No. 431, Part 1 (Boulder, 1980).
9. Chapman, S. & Bartels, J. *Geomagnetism* 1, 214 (1940).
10. Matsushita, S. in *Phys. Geomagn. Phenomena* 1, 301 (1967).
11. Richards, M. L. *Ann. Geophys.* 33, 177 (1977).
12. Wertheim, G. K. *Trans. Am. geophys. Union*, 35, 872 (1954).
13. Runcorn, S. K. *Nature* 202, 10 (1964).
14. Duffus, H. J. & Fowler, N. R. *Can. J. Earth Sci.* 11, 873 (1974).

Palaeomagnetic determination of emplacement temperature of Vesuvius AD 79 pyroclastic deposits

Dennis V. Kent*, Dragoslav Ninkovich*,
Tullio Pescatore† & Stephen R. J. Sparks‡

*Lamont-Doherty Geological Observatory, Palisades,
New York 10964, USA

†Institute of Geology and Geophysics, University of Naples, 80138,
Naples, Italy

‡Department of Earth Sciences, University of Cambridge,
Cambridge CB2 3EW, UK

The city of Herculaneum was buried under 20 m of pyroclastic deposits during the AD 79 eruption of Vesuvius, whose crater is only 7 km to the east. These deposits have been interpreted as the deposits of mudflows¹ or hot pyroclastic flows^{2–5}. Maury's studies⁶ of incinerated wood in Herculaneum demonstrate heating to at least 400 °C. We have studied the variation of remanent magnetism with temperature for specimens taken from the deposits, including specimens of the matrix material and of embedded lithic fragments. We conclude that the temperature of the deposit at emplacement is unlikely to have been greater than 400 °C, which further supports the interpretation of the pyroclastic deposits at Herculaneum as largely ignimbrites (hot pyroclastic flow deposits).

Ten oriented drill-core samples were collected from an excavation wall in the pyroclastic flow deposits at the western end of Herculaneum. The unit sampled is typical of the deposits

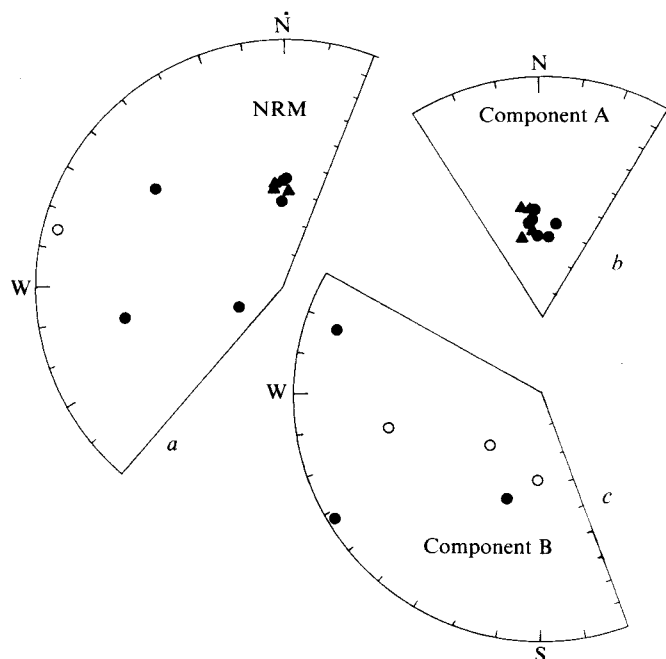


Fig. 1. Palaeomagnetic directions of pyroclastic samples from Herculaneum plotted on segments of equal-area projections. *a*, NRM directions; *b*, A components in lithic clast samples (●) and 400 °C thermally demagnetized NRM of matrix samples (▲); *c*, B components in lithic clasts (closed symbols on lower, open symbols on upper hemisphere of projection). Present geomagnetic field direction at sampling locality is dec = 359°; inc = 55°.

which bury the town and appears homogeneous and unstratified, consisting of a poorly sorted consolidated mixture of ash and pumice with local occurrences of angular lithic fragments of older volcanic rocks, limestone and brick. The volcanological stratigraphy at Herculaneum is detailed in ref. 7. The ash and pumaceous matrix is represented by samples H1 to H4; a variety of lithic fragments in addition to a large piece of light-coloured pumice, all between 10 and 20 cm in largest dimension, are represented by samples H5 to H10 (Table 1). The samples were taken over a stratigraphic interval of ~2 m, starting ~1 m above the excavation floor.

Most of the directions of the sample natural remanent magnetization (NRM) (Fig. 1*a*) group near to the present geomagnetic field for the site locality but four of the six samples of lithic fragments give divergent and scattered directions. The NRM intensities and Königsberger values (ratio of remanent magnetization to magnetization induced in present geomagnetic field) show only small variation amongst the samples of matrix but are more variable for the samples of lithic fragments (Table 1). In all samples, however, the remanent magnetizations exceed the induced components by anything between a factor of 4 and a factor of 69.

Progressive thermal demagnetization experiments show that the NRM of the matrix samples consists of a single stable component similar in direction to the present orientation of the geomagnetic field and which persists up to demagnetization temperatures of 550 °C, when it disappears (Fig. 2*a*). In contrast, the NRM of all six lithic fragments is the result of two components of magnetization: A and B. The A component is identified in the initial stages of thermal demagnetization up to temperatures of 350–400 °C. The B component is the magnetization remaining in the final stages of thermal demagnetization, up to temperatures between 550 and 600 °C when the magnetization of the samples disappears (Fig. 2*b, c*). The directions of the A and B component (Fig. 1*b, c* and Table 1) in each lithic fragment sample are readily determined by regression analysis on the linear segments of the orthogonal vector end point demagnetization diagrams, as shown in Fig. 2. Where two specimens could be obtained for measurement from a drill-core sample of a lithic clast, both specimens provided similar magnetic characteristics.

The mean direction of the A component from the six lithic samples is dec = 359.2°, inc = 56.8°, $\alpha_{95} = 4.2^\circ$ and is statistically indistinguishable from the mean direction of the matrix samples (Table 1) or from the present geomagnetic field direction (dec = 359°, inc = 55°). Although the B components are well characterized and show good stability in individual lithic clasts, as a group they are randomly distributed at the 95% level of confidence (Table 1).

The lithic fragments evidently possessed a remanent magnetization before their incorporation in the pyroclastic deposit. This magnetization is now preserved as component B and may have originated as a thermal remanence (TRM) during initial cooling of the basalt and pumice, possibly a detrital remanence during initial deposition of the limestone, and either

Table 1 Palaeomagnetic measurements on pyroclastic samples from Herculaneum

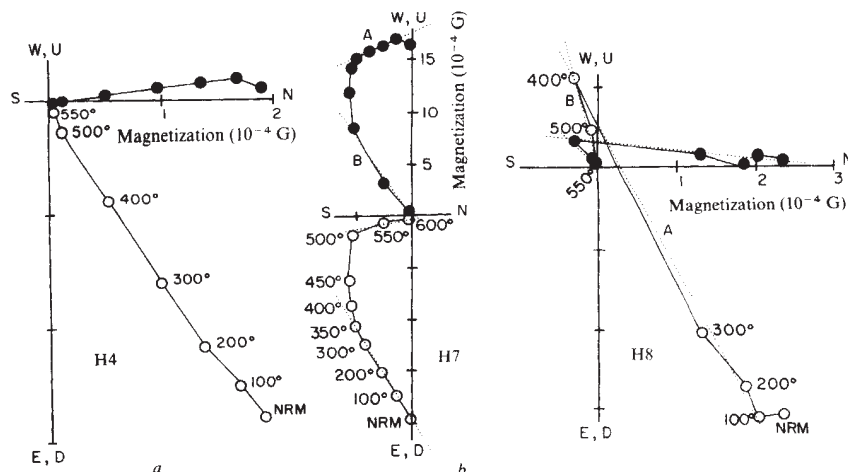
<i>a</i> Matrix samples		NRM				400 °C			
Sample		$J(10^{-4}\text{G})$	Q	dec	inc		dec	inc	
H1		3.4	5.8	355°	55°		351°	50°	
H2		3.2	5.9	353°	56°		349°	62°	
H3		3.1	5.2	1°	59°		355°	52°	
H4		3.4	5.6	357°	56°		355°	59°	
	Means ($n = 4$)			356.4°	56.5°		352.5°	55.8°	
				$(K = 1,020, \alpha_{95} = 2.9^\circ)$			$(K = 188, \alpha_{95} = 6.7^\circ)$		

<i>b</i> Lithic clasts		NRM				Component A		Component B	
Sample	Composition	$J(10^{-4}\text{G})$	Q	dec	inc	dec	inc	dec	inc
H5	Pumice	2.4	4.1	245°	73°	9°	56°	199°	53°
H6	Basalt	7.5	15	285°	−8	353°	55°	259°	−38°
H7	Basalt	10	14	259°	37°	356°	52°	238°	4°
H8	Limestone	3.5	69	358°	53°	6°	61°	226°	−66°
H9	Limestone	3.8	64	359°	61°	359°	61°	184°	−61°
H10	Brick(?)	15	23	309°	36°	354°	55°	286°	15°
	Means ($n = 6$)*			300.6°	49.9°	359.2°	56.8°	240.1°	−19.9°
				$(K = 4, \alpha_{95} = 37^\circ)$		$(K = 251, \alpha_{95} = 4.2^\circ)$		$(K = 2, \alpha_{95} = 59^\circ)$	
				$r = 4.80$		$r = 5.98$		$r = 3.75$	

J is intensity of magnetization; Q is Königsberger ratio; dec, inc, are declination, inclination of magnetization; means and statistics calculated using ref. 13; test for randomness from ref. 14.

*For $n = 6$, distribution is random if $r < 3.85$.

Fig. 2 Orthogonal projections of the successive end points of the magnetization vector (method from ref. 15) during progressive thermal demagnetization of representative samples from Herculaneum pyroclastic deposit. Solid symbols represent projection on the horizontal plane, open symbols on the north-south vertical plane. *a*, Matrix sample; *b*, basalt clast; *c*, limestone clast. Dashed lines in *b*, *c* are linear segments corresponding to A or B components.



a TRM or a chemical remanence when the brick was originally made. During incorporation in the pyroclastic deposit, physical rotation of the clasts resulted in the random distribution of their original (B) component magnetization.

After emplacement to their present position in the deposit, the lithic clasts became partially remagnetized, giving rise to component A which is statistically parallel in each clast. Given that the pyroclastic deposits probably experienced elevated temperatures, this magnetization is most likely a partial TRM—a supposition supported by the high Königsberger ratios. If a partial TRM mechanism is assumed for the A component, the magnetic unblocking temperatures measured in the laboratory can be used to approximate the maximum emplacement temperature from which the deposit cooled in place.

The curvature shown by some demagnetization trajectories (Fig. 2*b*) is interpreted as an overlap of blocking temperatures for magnetization components A and B. A conservative estimate for maximum unblocking temperature for component A would then correspond to the last point falling on the linear trajectory that defines this magnetization vector, which in Fig. 2*b* would be 350 °C. In other cases, (Fig. 2*c*), there is a strong inflection in the trajectories for components A and B and within the experimental uncertainties the maximum unblocking temperature for component A and the minimum unblocking temperature of the original component B coincide, at ~400 °C. The demagnetization behaviour of all lithic clast samples fell in either of these two categories and provided consistent unblocking temperature estimates for component A of between 350 and 400 °C.

This consistency over the various lithologies supports acquisition of a partial TRM at a common temperature and against elevation of blocking temperatures by other processes, for example, viscous magnetization or magnetochemical alteration. On the other hand, emplacement temperatures higher than 400 °C would have more completely reset the B component magnetization.

The fact that our maximum temperature estimate coincides closely with Maury's minimum temperature estimate strongly suggests that the emplacement temperature of the AD 79 Herculaneum pyroclastic deposit sampled was indeed close to 400 °C. The deposits can therefore be considered as largely ignimbrites (hot pyroclastic flows) which in general had not attained sufficiently high temperatures to cause pervasive welding. Other evidence for a high emplacement temperature includes the occurrence of vertical lithic-rich pipes, interpreted as degassing structures, and typical fine-grained ignimbrite basal layers to individual flow units²⁻⁵.

The matrix samples give only a single magnetization component which is similar in direction to the A component of the lithic clasts although the unblocking temperatures extend to 550–660 °C, ~200 °C higher than for component A. Given the close proximity of the sampling sites for the matrix and clast samples, it is unlikely that the higher maximum unblocking temperatures observed for the matrix material reflects a correspondingly higher emplacement temperature than indicated by the clast magnetizations. The unblocking temperature spectra of the matrix samples moreover are not consistent with the spectrum of a TRM produced by cooling a matrix sample from 600 °C in the Earth's field (Fig. 3). The NRM of the matrix material is therefore more likely to contain a contribution from a magnetization besides a partial TRM from 400 °C, perhaps a detrital remanence which, from the similarity in direction, was also acquired during emplacement of the pyroclastic deposit but which is typically carried by a grain assemblage with a wide distribution of blocking temperatures. Thus while bulk samples of the matrix may provide a valid record of geomagnetic field orientation corresponding to the time the deposits are formed, they are not as such readily useful for emplacement temperature determination.

Palaeomagnetic techniques have been used to discriminate hot and cold emplacement of pyroclastic deposits, for example, in Japan⁸, Santorini⁹ and the western US¹⁰. The Herculaneum pyroclastic flow can be categorized as a type III deposit (emplaced below their maximum blocking temperature but above the ambient temperatures) of Hoblitt and Kellogg¹⁰, of which they found no example. As palaeomagnetic techniques for emplacement temperature estimation depend largely on the common occurrence of magnetizable materials in pyroclastic deposits, the techniques should be of particular value in estimating emplacement temperatures of unwelded pyroclastic flow units in which lithological evidence for heating is ambiguous, as

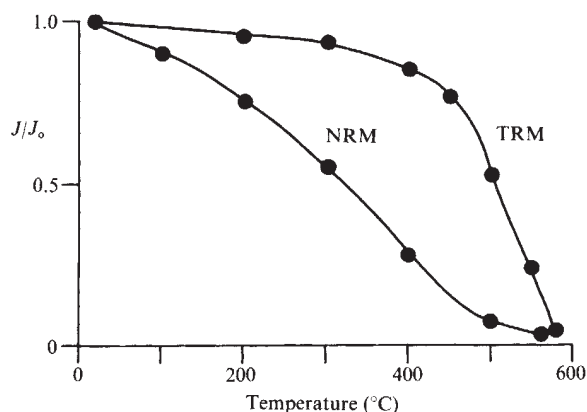


Fig. 3 Fraction of magnetization remaining with progressive thermal demagnetization for the NRM, and for a TRM produced by cooling from 600 °C in the Earth's magnetic field, of a sample of pyroclastic matrix from Herculaneum.

at Herculaneum. The size range of lithic fragments suitable for such experiments can be extended downward for example, to matrix components, by use of microanalytical techniques^{11,12} while unblocking temperature profiles through very large lithic clasts may provide additional information on the thermal history of the pyroclastic deposit. Moreover, estimates of maximum emplacement temperature by the magnetic method, which are possible up to the Curie point of the constituent magnetic mineral (for example, 585 °C for magnetite) can extend the range of temperature estimates obtained from studies of organic matter, which appear to be most definitive from 100 °C to ~350 °C (ref. 6).

We thank Doris Lafferty for laboratory magnetic measurements. Work was supported in part by NSF grant EAR79-09985 to D.V.K. and D.N. Lamont-Doherty Geological Observatory contribution 3114.

Received 3 November 1980; accepted 3 February 1981.

1. Maiuri, A. *Scient. Am.* **198**, 68 (1958); *Herculaneum* (Rome 1977).
2. Merrill, E. T. *Am. J. Archaeol.* **22**, 304 (1918); **24**, 262 (1920).
3. Sparks, R. S. J. & Walker, G. P. L. *Nature* **241**, 62 (1973).
4. Sparks, R. S. J., Self, S. & Walker, G. P. L. *Geology* **1**, 115 (1973).
5. Sparks, R. S. J. *New Scientist* **59**, 134 (1973).
6. Maury, R. C. *Bull. Centre Rech. Pan-SNPA* **10**, 289 (1976).
7. Sigurdsson, H., Cashdollar, S. & Sparks, R. S. J. *Am. J. Archaeol.* (in the press).
8. Aramaki, S. & Akimoto, S. I. *Am. J. Sci.* **255**, 619 (1957).
9. Wright, J. V. *Bull. volcan.* **41**, 131 (1978).
10. Hoblitt, R. P. & Kellogg, K. S. *Bull. geol. Soc. Am.* **90**, 633 (1979).
11. Wu, Y. T., Fuller M. & Schmidt, V. A. *Earth planet. Sci. Lett.* **23**, 275 (1974).
12. Van der Voo, R. & Geissman, J. W. *EOS* **59**, 1060 (1978).
13. Fisher, R. A. *Proc. R. Soc. A* **217**, 295 (1953).
14. Watson, G. S. *Mon. Not. R. ast. Soc., Geophys. Suppl.* **7**, 160 (1956).
15. Zijdeveld, J. D. A. in *Methods in Palaeomagnetism*, 254 (Elsevier, Amsterdam 1967).

Petrogenesis of Andean andesites from combined O-Sr isotope relationships

Russell S. Harmon

Isotope Geology Unit, Scottish Universities Research and Reactor Centre, East Kilbride G75 0QU, UK

Richard S. Thorpe & Peter W. Francis

Department of Earth Sciences, The Open University, Milton Keynes MK7 6AA, UK

The descent of the oceanic Nazca plate below the western margin of South America is responsible for the seismicity, tectonism and magmatism characteristic of Andean-type continental margins. Within this geological setting there is much debate concerning the petrogenesis of the calc-alkaline volcanic association. A thermal model for the descent of oceanic below continental lithosphere suggests that melting of subducted oceanic crust, the asthenospheric mantle wedge, and the lower crust might occur¹, and hence calc-alkaline volcanic rocks erupted at the surface in such regions might be derived from one or more of these sources. In the Andes, for example, Pichler and Zeil² have invoked partial melting of lower continental crust, while others³⁻⁵ envisage petrogenesis of andesites from sub-crustal mantle sources possibly in conjunction with crustal contamination⁶⁻⁹. We present here new O-isotope data for volcanic rocks from three contrasted volcanic settings in Ecuador, north Chile, and north-west Argentina. We then consider the significance of observed O-Sr isotope relationships, and attempt to evaluate the important factors in the petrogenesis of Andean andesitic lavas.

Active volcanism in the Andes occurs in three distinct regions: a northern zone (5°N–2°S), a central zone (16–28°S), and a southern zone (33–52°S); each having andesitic associations with distinctive chemical and isotopic compositions. The volcanic centres selected for study are in the northern and central volcanic zones. The northern zone in Ecuador and Columbia consists of composite volcanoes composed predominantly of basaltic andesite built on a 40–50-km thick sialic crust^{10,11}. Its sub-volcanic basement consists of metamorphic rocks of Cretaceous or older age¹². The central zone is characterized by composite volcanoes and large-volume ignimbrite sheets which overlie a thicker and more ancient sialic crust. Although the oldest rocks exposed adjacent to the active volcanic belt in southern Peru and northern Chile are Palaeozoic, Precambrian rocks occur to the west on the Peruvian coast¹³ and to the east below the Bolivian altiplano^{14,15} suggesting that the central volcanic belt is underlain by Precambrian rocks¹⁶. It has also been argued¹⁷ that the crust in the central Andes nearly doubled in thickness in the Cenozoic, so that the volcanoes of north Chile and south Peru rest on the central keel of the Andes where the crust is ~70 km thick. The volcanic setting in north-west Argentina is different to that to the west in southern Peru and northern Chile—here volcanic rocks occur on the 'Pampean Ranges' massif, a 'basin-and-range' province which comprises a wide variety of granitic and metamorphic rocks¹⁸ for which data indicate a late Precambrian to early Palaeozoic age¹⁹.

The Tertiary and Quaternary volcanic rocks of the Andes fall within the basaltic andesite–andesite–dacite classification of Peccerillo and Taylor²⁰. Chemical and isotopic compositions vary systematically across the Andean orogenic belt with SiO₂, K₂O and (⁸⁷Sr/⁸⁶Sr)_i increasing from west to east^{9,21-24} with calc-alkaline lavas generally enriched in K, Rb and Sr relative to P, Nb, Nd, Zr and Ti (refs 9, 25). There are, however, important inter-regional differences. The Ecuadorian lavas are olivine-pyroxene-bearing basaltic andesites with 53–61% SiO₂, 1.1–2.2% K₂O, 492–1,144 p.p.m. Sr, and 19–65 p.p.m. Rb (refs 7, 26, 27) whereas the lavas of north Chile and South Peru are predominantly pyroxene-hornblende-bearing andesites and

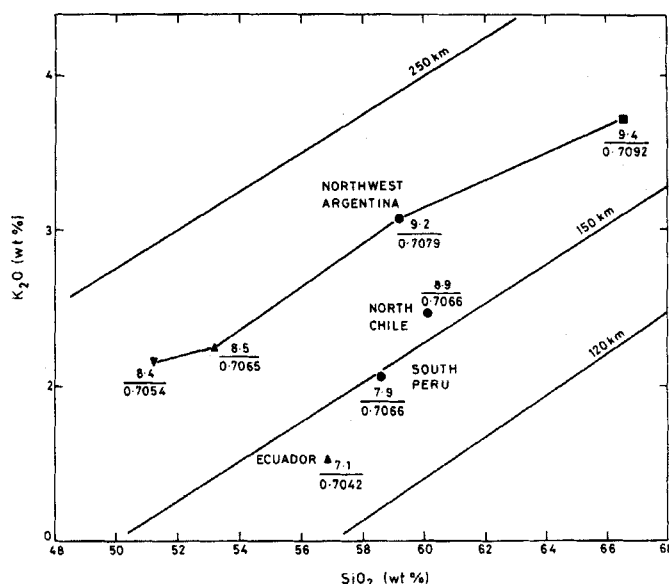


Fig. 1 Mean K₂O contents plotted against SiO₂ for volcanic centres in Ecuador, north Chile, south Peru, and north-west Argentina. The numbers shown by each data point are the average ¹⁸O/¹⁶O (upper) and ⁸⁷Sr/⁸⁶Sr (lower) ratios. The lines labelled 120, 150 and 250 km are inferred contours to the depth of the seismic zone³³. ▽, Basalt; ▲, basaltic andesite; ●, andesite; ■, dacite.

dacites with 56–66% SiO₂, 1.6–3.2% K₂O, 441–806 p.p.m. Sr, and 35–107 p.p.m. Rb (refs 17, 24). Thus, the respective mean SiO₂ and K₂O contents for the northern zone lavas are ~57 and 1.5% as compared with 60 and 2.2% for the western central zone lavas. By contrast, the Argentinean volcanic rocks cover the complete basalt–basaltic andesite–andesite–dacite spectrum. The basalt and basaltic andesites are olivine–pyroxene-bearing lavas largely from recent scoria cones and lava flows with 51–55% SiO₂, 2.0–2.7% K₂O, 477–1,048 p.p.m. Sr, and 45–75 p.p.m. Rb, whereas the andesites and dacites are pyroxene–hornblende-bearing lavas with 58–68% SiO₂, 2.9–4.2% K₂O, 236–645 p.p.m. Sr, and 52–269 p.p.m. Rb (ref. 28 and G. A. Kretschmar, personal communication).

¹⁸O/¹⁶O ratios were determined for 33 samples from the volcanic centres of Chimborazo, Tungurahua, Sincholhua, Cotopaxi, Punalica and Calpi in Ecuador, San Pedro and San Pablo in north Chile, and Cerro Galan in north-west Argentina (Fig. 1) using the same powders for which chemical composition as well as Sr isotope and Nd isotope ratios had previously been determined. ¹⁸O/¹⁶O ratios are presented here relative to SMOW, with pertinent chemical and isotopic data: overall analytical precision is 0.1–0.2‰. The combined O–Sr isotopic data are shown in Figs 1 and 2 with those of Magaritz *et al.*²⁹.

All samples analysed were petrographically fresh and there is no chemical or isotopic evidence to suggest that the lavas had been subject to significant hydrothermal or deuteric alteration. Water contents are low, varying from 0.8–0.55 wt % for basalts and basaltic andesites to 0.34–1.37 wt % for andesites and dacites. Low-temperature alteration products, veining and phenocryst reaction rims are absent. However, an isotopic study of the south Peruvian andesites²⁹ showed that petrographically fresh andesitic lavas had plagioclase $\delta^{18}\text{O}$ values which were lower by an average of 0.6‰ than the corresponding whole-rock $\delta^{18}\text{O}$ values; this was attributed to a small amount of post-eruptive hydrothermal alteration. In the Ecuadorian samples plagioclase $\delta^{18}\text{O}$ values are equivalent to whole-rock $\delta^{18}\text{O}$ values, reflecting crystal–liquid equilibrium at the time of eruption.

Combined isotopic studies have proved useful in identifying the source(s) and tracing the evolution of plutonic and volcanic rocks formed at convergent plate margins. The generation of calc-alkaline magmas is a complex process which might involve partial melting of subducted oceanic crust, the overlying asthenospheric mantle wedge, and/or lower continental crust as well as subsequent fractional crystallization and contamination by upper continental crust during transit. Here the combined O–Sr isotopic data for the Andean lavas are used to comment on the petrogenesis of andesites.

The Ecuadorian lavas are the least evolved of the Andean volcanics and have the lowest ¹⁸O/¹⁶O and ⁸⁷Sr/⁸⁶Sr ratios (6.5–7.7‰; 0.7038–0.7044). The small ¹⁸O- and ⁸⁷Sr-enrichment observed in these rocks relative to typical mantle (6.0 ± 0.5‰; 0.702–0.704) could result from direct involvement of slab-released volatiles in the partial melting of overlying mantle during dehydration of the slab, by partial melting of small amounts of hydrothermally altered oceanic slab itself followed by re-equilibration in the mantle, or by partial melting of a pre-existing, alkali metasomatized mantle. Although these stable isotopic data are inadequate for a choice to be made between these alternatives, the overall chemical and isotopic characteristics of the Ecuadorian lavas suggest an alkali-enriched mantle source.

By comparison with the Ecuadorian lavas, ¹⁸O/¹⁶O and ⁸⁷Sr/⁸⁶Sr ratios are higher and more variable for the central zone volcanics in north Chile (8.0–10.8‰; 0.7059–0.7072), south Peru (7.0–8.6‰; 0.7054–0.7076), and north-west Argentina (8.3–10.7‰; 0.7054–0.7097). Because fractional crystallization cannot significantly alter the O- or Sr-isotopic composition of a magma, the observed differences between the northern and central volcanic zones must indicate a different or additional factor in the petrogenesis of the central zone lavas. If the isotopic character of the Ecuadorian lavas is representative of pristine

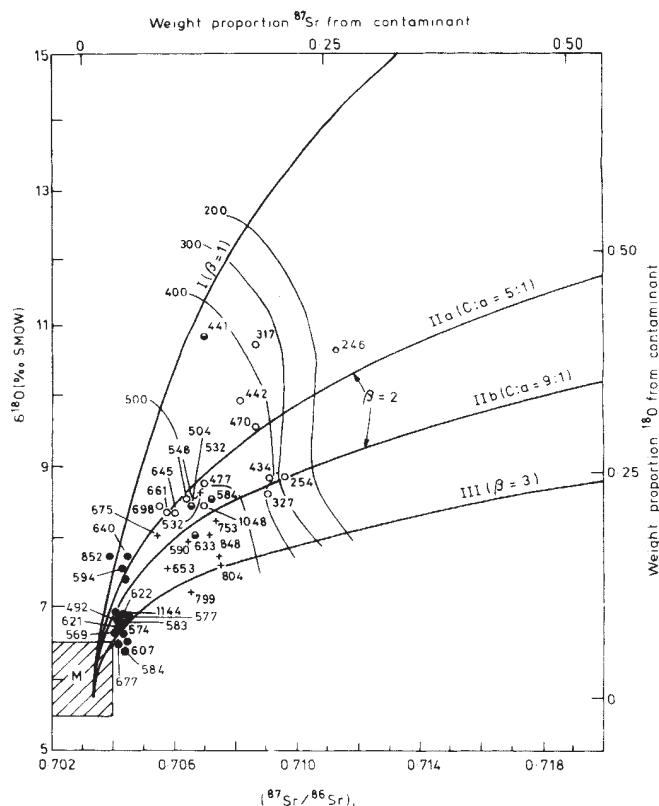


Fig. 2 Plot of $\delta^{18}\text{O}$ versus $(^{87}\text{Sr}/^{86}\text{Sr})$ for the Andean lavas analysed in this study. The solid curves I, IIa, IIb and III show the paths that would be followed during the process of combined assimilation-fractional crystallization for end-member compositions of hypothetical mantle ($^{18}\text{O} = 5.7\text{‰}$, $^{87}\text{Sr}/^{86}\text{Sr} = 0.703$) and an isotopically evolved crust ($^{18}\text{O} = 19\text{‰}$, $^{87}\text{Sr}/^{86}\text{Sr} = 0.735$). The curves, based on a ratio of Sr concentration in the end members of 5:1, are shown for three different (Sr cumulate/Sr melt) ratios (I, $\beta = 1$; II, $\beta = 2$ and III, $\beta = 3$) as well as two different ratios of cumulate to assimilated rocks (IIa, C:a = 9:1, IIb, C:a = 5:1) (for further discussion see ref. 34). ●, Ecuador; ○, North Chile; +, South Peru; ○, north-west Argentina. The number by each data point indicates the Sr content.

partial melts of a mantle source then those for the central zone lavas must reflect a different mantle or lower crustal source, or a contamination of magmas produced from the same mantle source.

Chemical and isotopic variations for the Andean lavas (Fig. 1) show that the K₂O–SiO₂ variations define linear trends which parallel those observed by Ninkovich and Hays³³ for island arcs and that for a given SiO₂ content there is a progressive increase in K₂O, ¹⁸O/¹⁶O and ⁸⁷Sr/⁸⁶Sr. Note that such increases occur in two directions on such a plot; enrichment of ¹⁸O and ⁸⁷Sr perpendicular to the depth contours of the seismic zone, reflecting inter-regional differences, and enrichment of ¹⁸O and ⁸⁷Sr parallel to depth contours indicating variation at a single volcano (for example, San Pedro–San Pablo, north Chile) or within a single volcanic province (for example, Cerro Galan, north-west Argentina). We interpret the former effect as indicative of source variation and the latter as contamination variation imposed on inherited source characteristics.

The most significant feature of the isotopic relationships for the Andean lavas is the positive correlation between ¹⁸O/¹⁶O and ⁸⁷Sr/⁸⁶Sr ratios (Fig. 1). Variations³⁴ of these two isotope ratios result from geologically independent processes, respectively the fractionation of stable O isotopes during the formation of low temperature sedimentary rocks at the Earth's surface and the radioactive decay of ⁸⁷Rb to ⁸⁷Sr. Thus a positive correlation

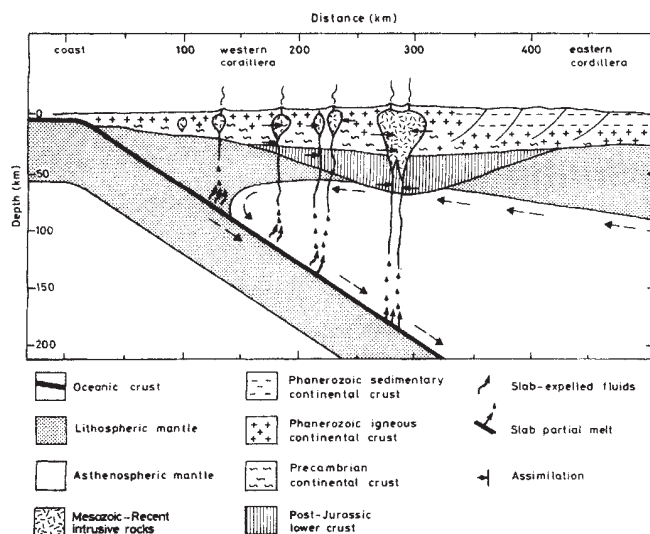


Fig. 3 The present crustal and mantle structure beneath the Andes at latitude 21°S; the known and inferred geological relationships; concepts regarding petrogenesis of Andean calc-alkaline magmas.

of ^{18}O and ^{87}Sr enrichment within a suite of genetically related igneous rocks must represent a major involvement with, or contamination by, continental crust. This correlation of O and Sr isotopic composition occurs within the Cerro Galan volcanic association of north-west Argentina, and also characterizes the overall O-Sr isotopic variation seen with the central and northern volcanic zones, although the relationship is different to those that have recently been described from two other areas of calc-alkaline magmatism, the Peninsular Range batholith of Southern California and Baja California³⁰ and the 'Newer' granites of the British Caledonides³¹. In each of these areas the positive correlations of $\delta^{18}\text{O}$ with $^{87}\text{Sr}/^{86}\text{Sr}$ extrapolate to O and Sr isotopic values representative of mantle compositions, suggesting that the calc-alkaline magmas in these three areas may have been largely derived from the mantle.

To evaluate the scale of the contamination variation, multi-component mixing models³⁴ have been used to evaluate the possible role of continental crust in the petrogenesis of the Andean lavas. Figure 2 shows the effects of a process of combined fractional crystallization-assimilation between a mantle and crystal end members with respectively low and high $^{18}\text{O}/^{16}\text{O}$ and $^{87}\text{Sr}/^{86}\text{Sr}$ ratios. Model lines are shown for ratios of the concentration of Sr in the cumulate material to that in the assimilated material (B) of 1 (curve I), 2 (curve II) and 3 (curve III). Curves II and IIB are respectively for bulk ratios of contaminant to assimilated material of 9:1 and 5:1. The positions of the curves in Fig. 2 depend on the exact compositions of the model end members. The low- ^{18}O , low- ^{87}Sr end member is reasonably well constrained within the limits $\text{Sr} = 800\text{--}1,000$ p.p.m.; $^{87}\text{Sr}/^{86}\text{Sr} = 0.702\text{--}0.704$, $^{18}\text{O}/^{16}\text{O} = 6.9 \pm 0.5$ considered characteristic of the mantle. By contrast, the composition of the high- ^{18}O , high- ^{87}Sr end member is arbitrary. The selected composition, $\text{Sr} = 50\text{--}200$ p.p.m.; $^{87}\text{Sr}/^{86}\text{Sr} = -0.735$, and $^{18}\text{O}/^{16}\text{O} = -19\%$ may be representative of the metamorphic and igneous rocks believed to be an important constituent of the central Andean Precambrian crust. However, the choice of a slightly different end-member composition would not change the shape of the curves, only the scale of the axes, and so require a larger or smaller proportion of crustal component to explain a given $^{18}\text{O}/^{16}\text{O}$ and $^{87}\text{Sr}/^{86}\text{Sr}$ value.

Figure 2 shows that the overall O-Sr isotopic trend for the Andean volcanic arc may be explained in terms of a process of combined assimilation and fractional crystallization between a

primitive, mantle-derived magma and isotopically evolved continental crust. This is best illustrated for the Cerro Galan volcanic suite of north-west Argentina. Francis *et al.*²⁸ considered two models for the Sr isotopic relationships (closed system bulk contamination and open system fractional crystallization-selective contamination) and concluded that the Sr-isotope variations were best explained by a small degree of selective Sr contamination of a magma ranging from basic to acidic in composition as the result of fractional crystallization. Figure 2 shows that the variation in $^{18}\text{O}/^{16}\text{O}$ ratios, 8.3–10.7%, together with an increase in $^{87}\text{Sr}/^{86}\text{Sr}$ ratio, 0.7054–0.7097, is consistent with a model involving fractional crystallization of a magma with the O- and Sr-isotopic characteristics of the least evolved sample (G49: $^{18}\text{O}/^{16}\text{O} = 8.4\%$, $^{87}\text{Sr}/^{86}\text{Sr} = 0.7054$) allied with up to 20% bulk assimilation. Although a well-defined O-Sr isotopic correlation is not present within the north Chilean and south Peruvian volcanics when considered individually, we feel this is likely to be due to the restricted compositional range of these lavas. However, it is possible that these variations might reflect primary characteristics of the source³², inherited characteristics of a source modified by crustal interaction³⁴ or the mixing of mantle and crustal magmas³⁵. However, the concomitant enrichment of ^{18}O and ^{87}Sr from the northern to central volcanic zones strongly indicates a crustal component to the central zone lavas.

As predicted by the assimilation fractional crystallization model, there is a general change in the major element chemistry coincident with the O-Sr isotopic variations (Fig. 1) as well as a general decrease in Sr content with increasing ^{18}O and ^{87}Sr content (Fig. 2). Figure 2 shows that most analyses cluster about the curves corresponding to a $\text{Sr}_{(\text{cumulate})}:\text{Sr}_{(\text{melt})}$ ratio of 2:1 and a cumulate to assimilate ratio of between 5:1 and 9:1. Therefore, if the assimilation-fractional crystallization process suggested by the O-Sr isotopic variations is valid, then the modelling implies that a large amount of cumulate must have formed during the evolution of the andesites and dacites from their basic precursors. Such cumulates might form parts of the batholithic intrusions inferred to be present below the Andes or might be present at depth within the lower crust. This suggestion is consistent with the recent estimate of the relative volumes of intrusive to extrusive rocks in the central Andes of Thorpe *et al.*³⁶, based on volumetric data.

Magaritz *et al.*²⁹ have recognized the necessity for an ^{18}O -enriched crustal component in the generation of andesitic lavas as a result of their O-Sr isotopic study of the Arequipa and Barroso volcanics of south Peru. After considering possible sources and by analogy with the Banda Arc, they concluded that the only reasonable source for the high ^{18}O component was continentally derived sediment which had been subducted and become involved in partial melting at depth. This argument was largely based on the absence of the $^{18}\text{O}/^{16}\text{O}$ - $^{87}\text{Sr}/^{86}\text{Sr}$ correlation which we have shown here to exist overall across the Andes. Even disregarding physical²⁸ and chemical²⁵ arguments against the involvement of subducted sediment in Andean calc-alkaline magma genesis, the isotopic data do not support such an idea. Certainly on mass balance considerations it is not feasible, given any reasonable proportions of components, to melt partially subducted-geosynclinal sediments ($\delta^{18}\text{O} \sim 13\%$; $^{87}\text{Sr}/^{86}\text{Sr} = 0.709$) together with altered oceanic crust ($\delta^{18}\text{O} \sim 7.9\%$; $^{87}\text{Sr}/^{86}\text{Sr} = 0.709$) and mantle to produce the ^{18}O - and ^{87}Sr -enriched lavas that occur in the central volcanic zone ($\delta^{18}\text{O} \sim 8\text{--}9\%$; $^{87}\text{Sr}/^{86}\text{Sr} \leq 0.711$). Likewise, the recent Pb-isotope data of Barreiro and Tilton³⁷ on the Arequipa and Barroso volcanics which are characterized by low $^{206}\text{Pb}/^{204}\text{Pb}$ ratios argue against contamination by subducted sediments—a crustal source is favoured.

The Andes are a classic setting in which to study the petrogenesis of calc-alkaline magmatism at destructive plate margins. Here the importance of subduction zone, mantle and crustal processes can be evaluated by chemical and isotopic techniques. The present data suggest an origin for andesitic magmas, illustrated in Fig. 3, which involves dehydration of the subducted

oceanic crust at depths of 60–100 km and alkali enrichment of the overlying asthenospheric mantle by LIL-enriched volatiles. The partial melting of this mantle then produces a basaltic melt slightly enriched in K, Rb and Sr relative to Zr and Ti and the REE as well as enriched in $^{87}\text{Sr}/^{86}\text{Sr}$ relative to $^{143}\text{Nd}/^{144}\text{Nd}$ and enriched in ^{18}O relative to normal mantle. This may be viewed as either a single-stage process in which alkali element enrichment occurs coincident with partial melting of primitive, homogeneous mantle, or as a two-stage process where a previously metasomatized mantle is partially melted to produce a calc-alkaline magma. Subsequent segregation of and diapiric uprise into the crust are followed by fractional crystallization and assimilation within the crust before eruption at the Earth's surface. In areas of relatively thin crust, such as in Ecuador, magmas rise to the surface as basaltic andesites largely unmodified by crustal contamination. However, in areas of thicker crust, such as in the central volcanic zone, the rising magmas are modified by a process of fractional crystallization allied with assimilation of isotopically evolved continental crust to produce andesites and dacites enriched in both ^{18}O and ^{87}Sr . Away from the subduction zone where partial melting occurs at depth in excess of 150 km, such as in north-west Argentina, magmas are probably derived from partial melting of the transformed oceanic slab with little or no volatile component so that the initial melts are enriched in both ^{18}O and ^{87}Sr relative to those magmas produced nearer the subduction zone from partial melting of the mantle. Here again slow rise of the basic melt through the thick crustal keel of the Central Andes provides an extended opportunity for fractional crystallization and crustal assimilation.

Isotopic work at the Scottish Universities Research and Reactor Centre and field work in the Andes by the Open University is supported by grants from the NERC. We thank G. A. Kretzschmar for unpublished K_2O data for Cerro Galan, J. Borthwick and J. Jocelyn for technical assistance, and C. J. Hawkesworth, H. P. Schwarcz and A. N. Halliday for helpful discussions. Miss J. McLean, Mrs E. Yuill, Mrs J. Bennett and D. MacLean helped to prepare the manuscript.

Received 26 August 1980; accepted 3 February 1981.

- Anderson, R. N., Delong, S. E. & Schwarz, W. M. *J. Geol.* **88**, 445–451 (1980).
- Pichler, W. & Zeil, W. *Bull. volcan.* **35**, 424–452 (1972).
- Noble, D. C., et al. *Geology* **3**, 501–504 (1975).
- James, D. E., Books, C. & Cuyubamba, A. *Bull. geol. Soc. Am.* **87**, 592–600 (1976).
- Brooks, C., James, D. E. & Hart, S. R. *Science* **193**, 1086–1089 (1976).
- Thorpe, R. S., Potts, P. J. & Francis, P. W. *Contr. Miner. Petrol.* **54**, 65–78 (1976).
- Francis, P. W., Moorbath, S. & Thorpe, R. S. *Earth planet. Sci. Lett.* **37**, 197–202 (1977).
- Dostal, J., Zentilli, M., Caelles, J. C. & Clark, A. H. *Contr. Miner. Petrol.* **63**, 113–128 (1977).
- Hawkesworth, C. J. et al. *Earth planet. Sci. Lett.* **42**, 45–57 (1979).
- Case, J. E., Barnes, J., Gabriel Paris, Q., Gonzalez, I. & Vina, A. *Bull. geol. Soc. Am.* **84**, 2895–2904 (1973).
- Meissner, R. O., Fleuh, E. R., Stibane, F. & Berg, E. *Tectonophysics* **35**, 115–136 (1976).
- Henderson, W. G. *J. geol. Soc. Lond.* **136**, 367–378 (1979).
- Shackleton, R. M., Ries, A. C., Coward, M. P. & Cobbold, P. R. *J. geol. Soc. Lond.* **136**, 195–214 (1979).
- Lehmann, A. *Geol. Rdsch* **67**, 270–285 (1978).
- Cobbing, E. J., Ozard, J. M. & Snelling, N. J. *Bull. geol. Soc. Am.* **88**, 241–246 (1977).
- Cobbing, E. J. & Pitcher, W. S. *Nature* **246**, 51–53 (1972).
- James, D. E. *Bull. geol. Soc. Am.* **82**, 3325–3346 (1971).
- Rapela, G. W. & Shaw, D. M. *Geochim. cosmochim. Acta* **43**, 1117–1129 (1979).
- Halpern, M. & Lattore, C. O. *Rev. Ass. Geol. Argentina* **28**, 195–205 (1973).
- Peccerillo, A. & Taylor, S. R. *Contr. Miner. Petrol.* **58**, 63–81 (1976).
- McNutt, R. H. et al. *Earth planet. Sci. Lett.* **27**, 305–323 (1975).
- Lefevre, C. *Contr. Miner. Petrol.* **41**, 259–272 (1973).
- Dupuy, C. & Lefevre, C. *Contr. Miner. Petrol.* **46**, 147–157 (1974).
- Francis, P. W., Robbol, M. J., Walker, G. P. L., Cobbold, P. R. & Coward, M. *Geol. Rdsch* **63**, 357–388 (1974).
- Klerks, J., Deutsch, S., Pichler, W. & Zeil, W. *J. Volcan. Geotherm. Res.* **2**, 49–71 (1977).
- Pichler, W., Stibane, F. R. & Weyl, R. *Neues Jb. Geol. Paläontol.* **2**, 102–126 (1977).
- Thorpe, R. S. & Francis, P. W. *Tectonophysics* **57**, 53–70 (1979).
- Francis, P. W., Thorpe, R. S., Moorbath, S., Kretzschmar, G. A. & Hammill, M. *Earth planet. Sci. Lett.* **48**, 257–267 (1980).
- Magaritz, M., Whitford, D. J. & James, D. E. *Earth planet. Sci. Lett.* **40**, 220–230 (1978).
- Taylor, H. P. & Silver, L. T. *U.S. geol. Surv. Open-File Rep.* **78**–701, 423–426 (1978).
- Harmon, R. S. & Halliday, A. N. *Nature* **283**, 21–25 (1980).
- Hamilton, P. J., O'Nions, R. K. & Pankhurst, R. J. *Nature* **287**, 279–284 (1980).
- Ninkovich, D. & Hays, J. D. *Earth planet. Sci. Lett.* **16**, 331–345 (1975).
- Taylor, H. P. *Earth planet. Sci. Lett.* **47**, 243–254 (1980).
- Halliday, A. N., Stephens, W. E. & Harmon, R. S. *J. geol. Soc. Lond.* **137**, 329–348 (1980).
- Francis, P. W., Thorpe, R. S. & Harmon, R. S. *Proc. R. Soc. Meet. on Origin and Evolution of the Earth's Crust* (Royal Society, London, in the press).
- Barreiro, B. & Tilton, G. R. *Proc. 26th int. geol. Congr.* **1**, 14 (1980).

Thin-skinned tectonics in the northern Rhenish Massif, Germany

R. Meissner & H. Bartelsen

Institut für Geophysik, Universität Kiel, 23 Kiel, FRG

H. Murawski

Geol.-Paläontol. Institut, Universität Frankfurt, 6 Frankfurt/Main, FRG

A network of profiles in the Hercynian mountain system (Geotransverse Rhenoharzynikum) is needed to determine details of the geological structure and find the reasons for tectonic movements. Outstanding features within the mountain system such as the large fault zone along the southern boundary of the Hunsrück¹ and the left flank of the Rhinegraben have already been investigated by reflection seismology^{2,3}. We describe here a new reflection–refraction experiment carried out in 1978 in the most northern part of the Rhenish Massif, across the Stavelot–Venn Massif near Aachen (Fig. 1). While layers in the lower crust dip to the NNW, a strong upper reflector at 3–4 km depth shows a slight dip to the SSE and indicates a prominent fault zone near Aachen, known in Belgium as the Faille du Midi. Here, a comparable and apparently identical reflector was found some 100 and 175 km to the west of our profile. We interpreted this reflector as a prominent and well lubricated thrust fault along which a huge horizontal nappe displacement took place during the last stages of the Variscan orogeny.

A special arrangement of shot points and geophone patterns was used to investigate the complicated tectonics in the northern part of the Rhenish Massif. Each of the 15 shots for the 31-km long line was recorded by three synchronized 48-channel recording units, each set out at 7.5-km length spreads, leaving two gaps of ~4 km each which were filled by individual reflection–refraction stations (see Fig. 2). Explosive charges of 50–400 kg were fired at points along the line and at their extension to the SSE, providing a total length of 37 km in the subsurface with a six-fold coverage for the main part of the profile. This simultaneous recording of such a long geophone spread provided us with huge move-out times, that is with the possibility of determining the velocity–depth structure with a very high accuracy from the reflection move-out. It was also a powerful method of determining the (predominantly horizontal) velocity from the evaluation of the first (refracted) arrivals which appeared as 12 overlapping refraction branches for the sixfold reflection coverage, hence providing sufficient redundancy for calculating near-surface velocities and corrections. (Details of this method and the interpretation work will be given elsewhere.) Figure 3 shows a stacked section of the upper 3 s two-way travel time involving the upper reflector. Strongly varying velocities within the steeply dipping surface layers are responsible for some remaining apparent undulations of the reflector. To determine its nature the following conclusions can be made.

The strong, high-frequency reflector indicates a considerable jump in impedance and hence in seismic velocity. The velocity cannot jump to a thick layer with a higher velocity because the first arrivals do not show any considerable velocity increase up to distances of 45 km, nor can the velocity jump to a thick layer with a lower velocity because no break in the first arrivals is observed. Hence, the strong reflectivity must result from a very thin layer of the order of a fraction of the dominant wavelength. Its thickness, estimated to be between 40 and 100 m, agrees both with the continuity of the first arrivals and with the single-wavelet, good quality of the reflector. A thin layer of higher velocity seems improbable because at least some high-frequency

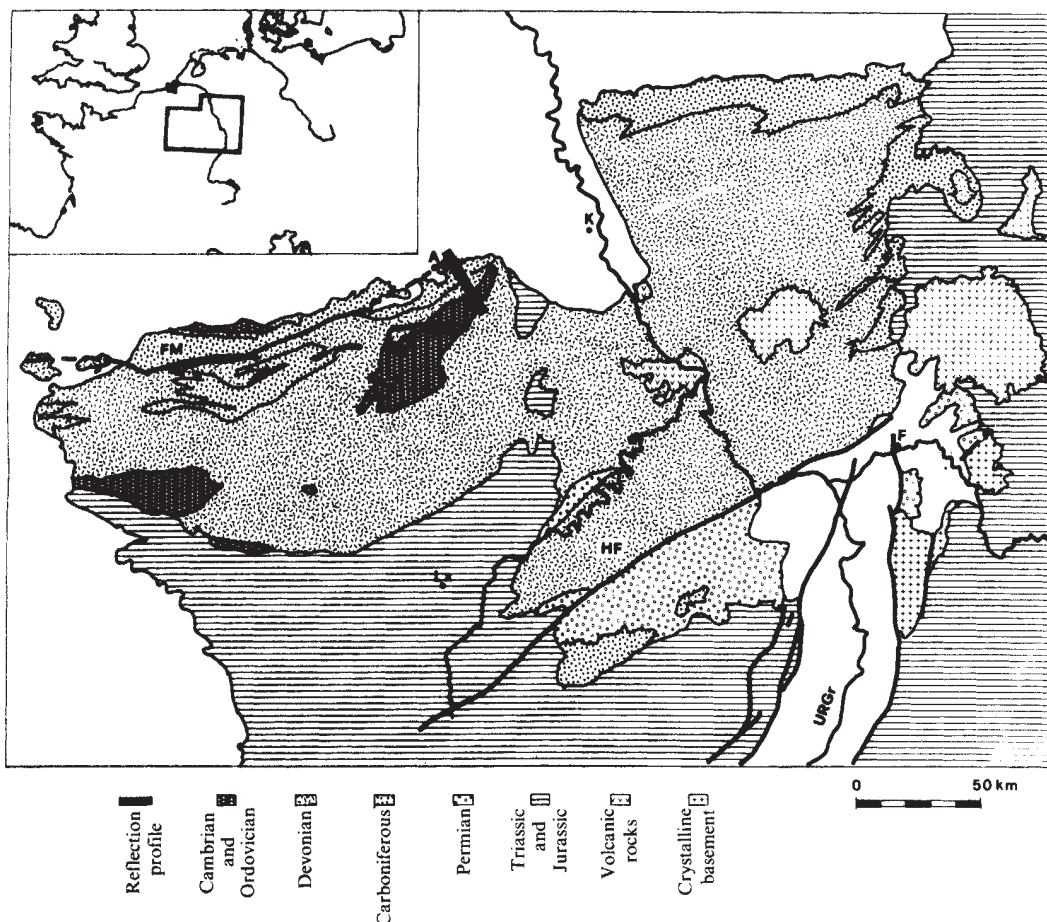


Fig. 1 Geological map of the Rhenish Massif and surrounding areas with location of profile. A, Aachen; F, Frankfurt; K, Köln; Lx, Luxembourg; FM, Faille du Midi; URGr, Upper Rhine graben; HF, Hunsrück fault.

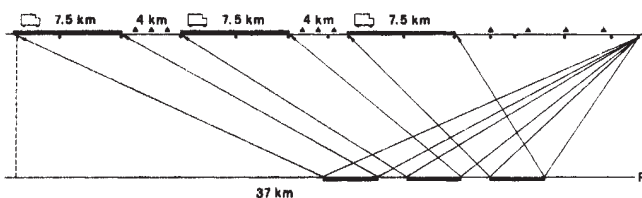


Fig. 2 Profile field set-up. ●, Shotpoints; ▲, additional wide-angle stations. Hatching on the schematic reflector R shows the 'subsurface coverage' for the first shotpoint, corresponding to the geophone spread of three recording units. Signals from all shotpoints were recorded by the geophone spreads resulting in a sixfold coverage of reflectors.

events with higher velocities would be expected in this case. The only consistent solution, therefore, is provided by a thin layer with lower velocities. Such a layer, together with the shape and the direction of the reflector which points directly towards a prominent surface fault system in the Hohes Venn area, strongly suggests that the reflecting interface resembles a thrust fault with

indications of large nappe displacement to the north-west. Figure 4 shows a geological cross-section of the upper part of the area under investigation, containing the reflector as its lower boundary.

The interpretation of the reflector as a prominent thrust fault is supported by petrological evidence, seismological and geodetic observations and by other seismic reflection results. Thrust faults, for example, generally exhibit high pore pressure, grain diminution and generally, a hydrolytic weakening⁴. Although these effects lead to a decrease in velocities^{5,6}, a high pore pressure which considerably reduces the effective pressure might have the strongest effect on the velocity decrease within a thin zone of deformation⁷. Hence it seems important that the fault zone is still seismically active in the correct overthrust direction as seen from a fault plane solution of a quake in the thrust fault area⁸. Also, an increased vertical uplift has been observed in the Hohes Venn area⁹. Today's stress pattern¹⁰ seems to be similar to that during late Variscan time, hence probably reactivating the old fault zones. Note that reflection experiments of the Famenne area, ~100 km WSW of the

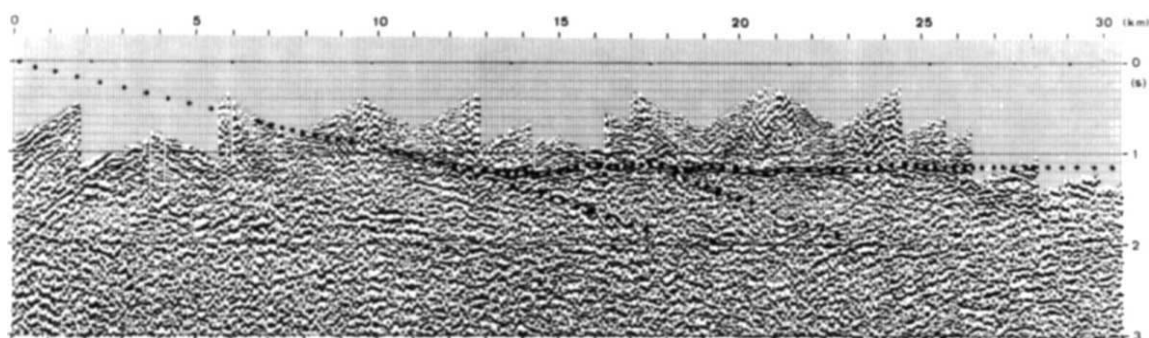


Fig. 3 Stacked section of the upper 3 s with reflector at 1.2 s two-way travel time.

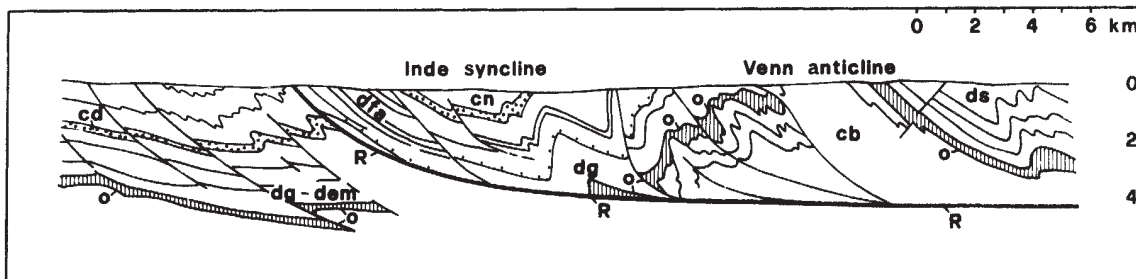


Fig. 4 Geological cross-section along the profile. R, Seismic reflector; cb, Cambrian; o, Ordovician. Devonian: dg, Gedinnian; ds, Siegenian; dem, Emsian; dfa, Famennian. Carboniferous: cd, Dinantian; cn, Namurian.

present profile, have also revealed a prominent reflector at roughly the same depth as our presumed thrust fault and also dipping gently to the SSE. A comparable reflector has also been found some 175 km to the west¹² and even farther. Figure 5 shows the approximate depth of this reflector in the area of investigation. The identified reflector, interpreted as a thrust fault, covers a large area and may even extend laterally to the east and west and to greater depths to the south where no seismic investigation has been carried out.

Recent investigations in the southern Appalachians by the COCORP group¹³ and in the northern Appalachians have shown similar pronounced, shallow and slightly inclined reflectors. In the southern part they are interpreted as thrust faults on which the crystalline basement of the Appalachians has been thrust at least 260 km towards the North American continent¹³⁻¹⁵.

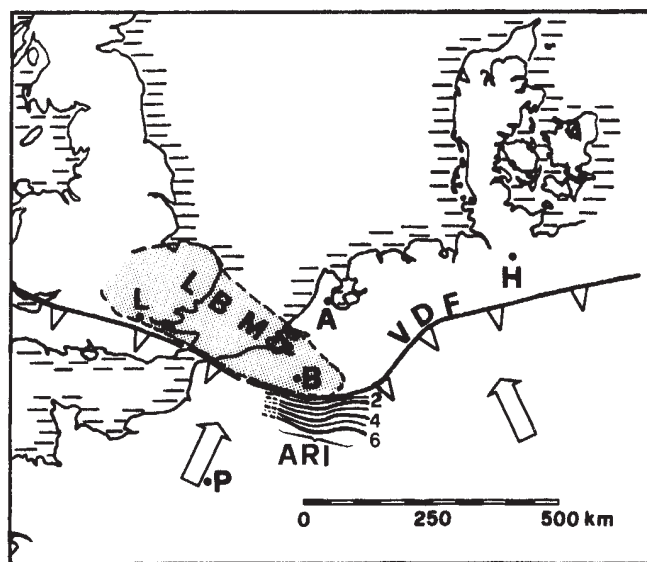


Fig. 5 Area of reflection investigation (ARI) within the Variscan deformation front (VDF). ARI shows isolines of depth of reflecting interface (thrust fault) in km. LBM, London Brabant massif; H, Hamburg; L, London; A, Amsterdam; B, Bruxelles; P, Paris.

We do not know whether the thin-skinned nappe in our test area in the Variscides was transported over a comparably large distance, but this possibility must be considered. The lateral dimensions of the nappe of at least 200 km—a lower limit, only marked by the available seismic reflection observations—indicates that large and regional compressive forces were acting during the last phases of the Variscan orogeny. We also do not know whether a pattern of approaching and colliding island arcs and microcontinents, as apparently involved in the creation of the southern Appalachians¹⁵, was also responsible for the formation of the complex Variscan tectonics. Certainly, here the deformation front also moved from SW towards the NW—towards a large continental area.

Extensive nappe displacements are well known from virtually all orogenic belts from Caledonian to recent compressional regimes, for example, Alpine-Himalayan; but in most cases the depth, the dimension and the structure of them remain unknown. For parts of the Appalachians^{14,15}, and now also for parts of the northwestern Variscides the character and the depths of the thrust planes have been revealed and their nature analysed from reflection seismology. The formation of thin-skinned nappes riding over a rather undeformed subsurface along plane, well lubricated thrust faults may indicate a final stage of compressional tectonics.

We thank P. Ziegler, for helpful discussions and our colleagues of the Geological Institute of the RTHW Aachen for the geological profile.

Received 14 October 1980; accepted 3 February 1981.

1. Ziegler, P. A. *Geologie Mijnb.* **57**, 589–626 (1978).
2. Ahorner, L. & Murawski H. Z. *ges. Geol.* **126**, 63–82 (1975).
3. Meissner, R., Bartelsen, H. & Murawski, H. *Tectonophysics* **64**, 59–84 (1980).
4. Etheridge, M. A. & Wilkie, J. C. *Tectonophysics* **58**, 159–178 (1979).
5. Kern, H. *Tectonophysics* **44**, 185–203 (1978).
6. Meissner, R. & Fakimi, M. *Geophys. J. R. astr. Soc.* **49**, 133–143 (1977).
7. Christensen, N. J. *J. Geophys. Res.* **84**, 6849–6858 (1979).
8. Ahorner, L. *Protokoll 3. DFG Kolloq.*, Neustadt 40–42 (1978).
9. Teichmüller M. & Teichmüller, R. *Fortschr. Geol. Rheinld Westf.* **27**, 323–355 (1979).
10. Illies, J. H. & Greiner, G. *Tectonophysics* **52**, 349–359 (1979).
11. Bless, M. J. M., Bouckaert, J. & Paproth, E. *Meded. Rijks geol. Dienst* **32-1**, 3–13 (1980).
12. Clément, J. *Ann. Soc. géol. Nord* **83**, 237–241 (1963).
13. Brewer, J. A. & Oliver, J. E. *A. Rev. Earth planet. Sci.* **8**, 205–230 (1980).
14. Cook, F. A. *et al. Geology* **7**, 563–567 (1979).
15. Cook, F. A., Brown, L. D. & Oliver, J. E. *Scient. Am.* **243**, 124–138 (1980).

Group selection is implicated in the evolution of female-biased sex ratios

Robert K. Colwell

Department of Zoology, University of California, Berkeley, California 94720, USA

Female-biased sex ratios are common among arthropods that mate in small groups with episodic dispersal¹⁻³. When breeding groups are founded by more than one female, individual (within-group) selection favours an unbiased (1:1) sex ratio, regardless of group size and degree of inbreeding (or sibmating), and regardless of the intensity of local competition among siblings for mates. On the other hand, female-biased sex ratios are favoured by group selection in randomly structured populations. Existing treatments of the evolution of female-biased sex ratios^{1,2,4-6} correctly predict the result of evolution, but do not make clear that it is a balance between opposing forces at two levels of selection³. I now show this to be so.

Fisher⁷ pointed out that in a random-mating population with a biased sex ratio, an individual of the rarer sex leaves more progeny, on average, than an individual of the commoner sex, so that frequency-dependent selection favours equal investment in sons and daughters. Hamilton¹ modelled the evolution of

Table 1 Model for the evolution of female-biased sex ratios as a balance between the opposing forces of individual and group selection

For all breeding groups, let

- n = number of inseminated females founding each group
- a, b = proportion of sons among the progeny of each A-type (a) or B-type (b) female
- r = number of progeny produced by each female, regardless of type
- $\bar{p}$ = proportion of founding females that are A-type in the global population
- p' = proportion of grandmaternal gene copies that are due to A-type founding females in the global population

I Individual selection (within groups)

In a group founded by n inseminated females, of which $x = p_x n$ are A-type (each producing ar sons) and $(1 - p_x)n$ are B-type (each producing br sons), the numbers of sons and daughters among the group's progeny (F_1) are

$$s_x = n[p_x a + (1 - p_x)b] \quad \text{and} \quad d_x = n[p_x(1 - a) + (1 - p_x)(1 - b)] \quad (1)$$

Now the F_1 mate at random within the group, with mating probabilities

$$\begin{aligned} A\sigma \times A\sigma: M_{AA} &= p_x^2 \left(\frac{a}{s_x} \right) \left(\frac{1-a}{d_x} \right) r^2 n^2 \\ A\sigma \times B\sigma: M_{AB} &= p_x(1-p_x) \left(\frac{a}{s_x} \right) \left(\frac{1-b}{d_x} \right) r^2 n^2 \\ B\sigma \times A\sigma: M_{BA} &= (1-p_x)p_x \left(\frac{b}{s_x} \right) \left(\frac{1-a}{d_x} \right) r^2 n^2 \\ B\sigma \times B\sigma: M_{BB} &= (1-p_x)^2 \left(\frac{b}{s_x} \right) \left(\frac{1-b}{d_x} \right) r^2 n^2 \end{aligned} \quad (2)$$

$$M_{AA} + M_{AB} + M_{BA} + M_{BB} = 1$$

where $A\sigma$ means a son of an A-type founder, $B\sigma$ a daughter of a B-type founder, and so on. The proportion of grandmaternal gene copies among the F_2 that are due to A-type founders is

$$p'_x = M_{AA} + \frac{1}{2}(M_{AB} + M_{BA}) = \frac{p_x^2 a(1-a) + \frac{1}{2}p_x(1-p_x)[a(1-b) + b(1-a)]}{[p_x a + (1-p_x)b][p_x(1-a) + (1-p_x)(1-b)]} \quad (3)$$

and the total number of F_2 gene copies due to A-type founders is

$$p'_x V_x = p'_x r d_x \quad (4)$$

assuming diploidy.

For a particular breeding group, let

- x = number of founding females that are A-type (the rest, $[n - x]$, are B-type)
- $p_x = x/n$ = proportion founding females that are A-type
- p'_x = proportion of grandmaternal gene copies that are due to A-type founding females
- s_x, d_x = total number of sons (s_x) or of daughters (d_x) produced by all n founding females
- $V_x = r d_x$ = total number of grandprogeny (F_2) produced by the daughters of all n founding females (group productivity)
- π_x = probability that a particular group is founded by x A-type and $(n - x)$ B-type females

II Integration of individual selection with group selection

In the global population, the distribution of group composition is assumed to be binomial, so the probability that a group will have x A-type females among its n founders is

$$\pi_x = \binom{n}{x} \bar{p}^x (1 - \bar{p})^{n-x} \quad (5)$$

The average number of F_2 gene copies, per group, due to all founding females is

$$\bar{V} = \sum_{x=0}^n \pi_x V_x = n r^2 [\bar{p}(1-a) + (1-\bar{p})(1-b)] \quad (6)$$

since $E[p_x] = \bar{p}$. The average number of F_2 gene copies, per group, due to A-type founding females is

$$\sum_{x=0}^n \pi_x p'_x V_x \quad (7)$$

Therefore, the proportion of grandmaternal gene copies in the F_2 that are due to A-type founding females in the global population is

$$\bar{p}' = \sum_{x=0}^n \pi_x p'_x V_x / \bar{V} \quad (8)$$

For example, for $n = 2$,

$$\bar{p}' = \frac{\bar{p}(1-\bar{p})[(2a+b)(1-a) + a(1-b)]}{2(a+b)} + \bar{p}^2(1-a) \quad (9)$$

female-biased sex ratios for a population structured into randomly initiated groups, each founded by n inseminated females, whose offspring mate locally among themselves and then disperse to found new groups. For such a population, Hamilton showed that an autosomal gene that causes its female bearer to produce ar sons among her r progeny, where $a = (n-1)/2n$, will be 'unbeatable' by a gene coding for any other sex ratio (diploidy, complete mating and equal investment per offspring assumed). Thus, the optimal progeny sex ratio is $1/4$ sons with $n = 2$ founding females, rising asymptotically toward $1/2$ sons with increasing n . A solitary founding female ($n = 1$) should produce the minimum number of sons required to inseminate her daughters.

Hamilton¹ and others^{4-6,8} have suggested that daughter-biased progenies are an adaptive response to selective pressures on the mother, serving to reduce 'local mate competition' (LMC) among her sons. By this argument, the increasing probability that brothers will compete for the same mates, as the number of sibships becomes smaller, is countered by a progressive decrease in the proportion of sons. Members of a local breeding group, such as the wasps in a single fig² or host pupa⁶, or the hummingbird flower mites in a single inflorescence³ are highly inbred, relative to the global population of which they are a part. Hamilton's¹ early emphasis on inbreeding, and Maynard Smith's⁴ reformulation of Hamilton's LMC model in terms of the proportion of matings that are between siblings, may have suggested to some authors⁸⁻¹⁰ that inbreeding (or sibmating) itself somehow provides a mechanism for the evolution of

daughter-biased progenies as an individually advantageous adaptation, reducing sibling competition for mates.

For the case of more than one founding female ($n > 1$), a simple algebraic model for the evolution of progeny sex ratio (Table 1) indicates that (1) LMC itself cannot select for female-biased progenies, (2) inbreeding (sibmating) is not relevant to the problem, and (3) group selection (differential productivity of genetically different groups) is required for the evolution of female-biased sex ratios.

In the model, each generation of local mating among the progeny of n founding mothers is followed by population-wide mixing and dispersal of their inseminated daughters. Two maternal phenotypes ('A-type' and 'B-type') compete, each producing a different, fixed progeny sex ratio (a and b , respectively, as proportion of sons). The fitness of a founder is measured in terms of the gene copies she leaves among grandprogeny^{1,4,6,11}. Figure 1 shows that the fitness of a 'fisherian' founder, producing an unbiased progeny sex ratio ($b = 0.5$), always exceeds the fitness of a 'hamiltonian' founder, producing a daughter-biased sex ratio ($a < 0.5$), regardless of the number of founders, initial proportion of the two types and degree of female bias in hamiltonian progenies¹. Because the degree of selective advantage of fisherian females depends only on their proportion among founders, relative fitness of the two types is completely unaffected by the level of inbreeding or sibmating, and by the intensity of competition among brothers for mates, which are functions of the number of founders, given a and b . Individual selection on females whose offspring breed in local

groups with the offspring of other founders, however few, cannot favour daughter-biased progenies.

How, then, can female-biased sex ratios evolve? If the dispersal pool includes more than one sex-ratio phenotype, local groups will vary by chance in the frequency of the phenotypes among founders. Within each group, the relative fitness of fisherian founders exceeds that of hamiltonian founders. However, groups with higher-than-average initial frequencies of hamiltonian females produce a greater-than-average absolute number of grandprogeny, simply because their founders jointly produce more daughters (and fewer sons) than

Table 2 Individual and group selection in the evolution of female-biased sex ratios: a numerical example with $n = 4$

	Group 1	Group 2
I Initial composition of groups		
A. Number of hamiltonian ($a = \frac{3}{4}$) founders: x	3	1
B. Number of fisherian ($b = \frac{1}{4}$) founders: $(n - x)$	1	3
C. Proportion hamiltonian founders: p	0.75	0.25
II Contribution of founders to grandprogeny:		
Individual selection		
A. Proportion of grandmaternal gene copies due to		
1. Each hamiltonian founder	0.2470	0.2471
2. Each fisherian founder	0.2591	0.2510
3. All hamiltonian founders: p'_x	0.7409	0.2471
4. All fisherian founders: $(1 - p'_x)$	0.2591	0.7529
B. Fitness of a hamiltonian founder relative to a fisherian founder	0.9533	0.9845
III Differential productivity of groups:		
Group selection		
A. Total grandprogeny produced by each group: V_x	$2.375r^2$	$2.125r^2$
B. Number of grandprogeny gene copies within each group due to		
1. Each hamiltonian founder	$0.5865r^2$	$0.5251r^2$
2. Each fisherian founder	$0.6154r^2$	$0.5333r^2$
3. All hamiltonian founders: $p'_x V_x$	$1.7596r^2$	$0.5251r^2$
4. All fisherian founders: $(1 - p'_x) V_x$	$0.6154r^2$	$1.5999r^2$
IV Number of gene copies among pooled grandprogeny		
A. Due to each hamiltonian founder: $\frac{1}{4}(1.7596 + 0.5251)r^2 =$	$0.5712r^2$	
B. Due to each fisherian founder: $\frac{1}{4}(0.6154 + 1.5999)r^2 =$	$0.5538r^2$	

Individual and group selection in the evolution of female-biased sex ratios. This example is based on an arbitrary initial composition of groups. Two groups are formed by sampling from a founder pool that consists of half fisherian (F) and half hamiltonian (H) inseminated females. The groups differ in the frequency of the two types (line IC) but together reflect the (equal) frequencies in the founder pool. Each female in each group then produces r progeny, of which $1/2 r$ are sons, for each fisherian mother, and $3/8 r$ are sons, for each hamiltonian mother. The progeny mate at random among themselves within each group (all females are inseminated). Computing fitness within groups as the expected number of grandmaternal gene copies due to each founding female (lines IIIB, 1, 2), it is evident that the individual fitness of each fisherian female exceeds that of each hamiltonian female within groups (line IIB). This is true for any mixed group (Fig. 1). On the other hand, group fitness—the total number of grandprogeny produced (line IIIA)—is greater for the group with the higher initial frequency of hamiltonian females, because its founders jointly produce more daughters (and fewer sons) than do the predominantly fisherian founders of the other group. Because the more productive group is mostly hamiltonian, when the number of gene copies due to each of the two types of founders is evaluated among the pooled grandprogeny of the two groups, the hamiltonian founders prove to have contributed more than the fisherian founders (lines IVA, B), in spite of the selective disadvantage of hamiltonian founders within each group. The force of group selection depends on the variance in initial composition of groups, here arbitrarily established to make a point. In fact, random (binomial) variation is sufficient for the force of group selection to exceed that of individual selection^{3,13}. Greater-than-binomial variance only increases the effect^{12,13}. In terms of Price's^{2,14} hierarchical model of selection, the fitness of groups (V_x) co-varies positively with the proportion of hamiltonian founders that initiate them (p_x), whereas the fitness of individuals within groups co-varies negatively with possession of the hamiltonian trait (or, genetically, with the number of hamiltonian alleles an individual carries): p'_x is inevitably less than p_x . Thus, female-biased sex ratios are opposed by individual selection, but are favoured by group selection.

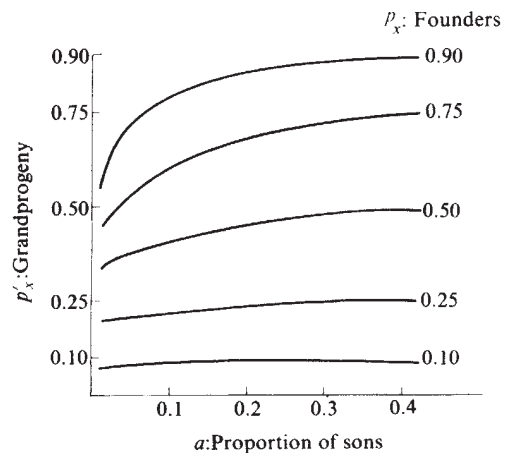


Fig. 1 Fisher's⁷ sex ratio principle. The selective disadvantage of producing female-biased progeny. A group is founded by hamiltonian females in proportion p_x , each producing proportion $a < 0.5$ sons, and by fisherian females in proportion $(1 - p_x)$, each producing half sons and half daughters. Here, the proportion of grandmaternal gene copies in the grandprogeny that are due to the hamiltonian founders, p'_x (computed by equation (3), Table 1) is plotted as a function of a for several values of p_x . The figure demonstrates that p'_x is inevitably less than p_x : the fitness of fisherian founders always exceeds that of hamiltonian founders in a mixed group. Note that the number of founding females (n , group size) plays no part in this result. Fisher's principle operates with equal force in random-mating groups or populations of any size, whether initiated

average. The crucial point is this: when genetic contributions to the grandprogeny of randomly initiated groups are pooled by dispersal, hamiltonian founders prove to have made a greater per capita contribution than fisherian founders, even though within every mixed group the reverse is true. Table 2 provides a numerical example of this phenomenon.

In general terms, the global proportion of grandmaternal gene copies in the grandprogeny that are due to A-type founders is a productivity-weighted average of local (within-group) proportions, derived in part II of Table 1 on the assumption of random (binomial) variation among groups in the initial frequency of A-type and B-type founders. When A-type females are optimally hamiltonian [$a = (n - 1)/2n$] it is easily shown numerically (for example, with equation (9), Table 1) that the global proportion of grandmaternal gene copies due to hamiltonian females is greater than the proportion of hamiltonian females among group founders, regardless of the progeny sex ratio produced by B-type founders and initial frequencies of the two types. (Hamilton¹ obtained this result analytically by direct optimization.) In other words, whatever the mode of inheritance, optimally hamiltonian alleles will increase in frequency relative to fisherian (or any other) alleles in a randomly structured population. The present model permits a more general statement: if $b < a \leq (n - 1)/2n$ or if $b > a \geq (n - 1)/2n$, A-type females prevail at any starting frequency; if $a < (n - 1)/2n < b$, a frequency-dependent stable polymorphism exists (for example, if $n = 3$, $a = 0.1$, and $b = 0.6$, then $\bar{p}' = \bar{p} = 0.48$).

The force of individual selection is independent of the number of founders (n) (Fig. 1), whereas the force of group selection increases with genetic variation among groups^{12,13}, which becomes greater as n decreases, producing increasingly female-biased sex ratios. If among-group variation in progeny sex ratio is suppressed, or if it is permitted, but differential productivity is eliminated by the imposition of a carrying capacity, female-biased sex ratios cannot evolve, regardless of n , and therefore regardless of the level of inbreeding, sibmating and local mate competition³.

I thank D. S. Wilson for consultations and encouragement, and H. J. Brockmann, E. L. Charnov, A. B. Clark, L. Goldwasser, M.-C. King, S. Levin, J. Maynard Smith, P. W. Sherman, N. C. Stenseth and J. H. Werren for stimulating discussions. This work was supported by the NSF (DEB78-12038).

Received 4 August 1980; accepted 19 January 1981.

- Hamilton, W. D. *Science* **156**, 477-488 (1967).
- Hamilton, W. D. in *Sexual Selection and Reproductive Competition in Insects* (eds Blum, M. S. & Blum, N. A.) 167-221 (Academic, New York, 1979).
- Wilson, D. S. & Colwell, R. K. *Evolution* (in the press).
- Maynard Smith, J. *The Evolution of Sex* (Cambridge University Press, 1978).
- Bulmer, M. G. & Taylor, P. D. *Nature* **284**, 448-449 (1980); *J. theor. Biol.* **86**, 83-89 (1980).
- Werren, J. H. *Science* **208**, 1157-1159 (1980).
- Fisher, R. A. *The Genetical Theory of Natural Selection* (Oxford University Press, 1930).
- Alexander, R. D. & Sherman, P. W. *Science* **196**, 494-500 (1977).
- Trivers, R. L. & Hare, H. *Science* **191**, 249-263 (1976).
- Stenseth, N. C. *Oikos* **30**, 83-89 (1978).
- Shaw, R. F. & Mohler, J. D. *Am. Nat.* **87**, 337-342 (1953).
- Slatkin, M. & Wade, M. J. *Proc. natn. Acad. Sci. U.S.A.* **75**, 3531-3534 (1978).
- Wilson, D. S. *The Natural Selection of Populations and Communities* (Benjamin/Cummings, Menlo Park, 1980).
- Price, G. A. *Ann. hum. Genet.* **35**, 485-490 (1972).

Growth hormone stimulates adrenal steroidogenesis in the fetus

Uday P. Devaskar*, Sherin U. Devaskar,
Sandra Voina, Nelson Velayo & Mark A. Sperling

Department of Pediatrics, Division of Neonatology and Endocrinology, Children's Hospital Research Foundation, University of Cincinnati Medical Center, Cincinnati, Ohio 45267, USA

The 'surge' of corticosteroid in fetal plasma during late gestation has been implicated in the initiation of parturition and the maturation of enzyme systems in organs such as the lung, liver, adrenal medulla and thyroid¹⁻⁵. But in all species studied, the mechanism responsible for increased secretion in the fetus remains unclear. The hormone adrenocorticotropin (ACTH) is well established as the primary regulator of adrenocortical cellular growth and secretory function during fetal and adult life⁶. However, no increase in fetal plasma ACTH has been observed before the corticosteroid surge in sheep or humans^{7,8}, although differential responsiveness of the fetal adrenal cortex to ACTH at various gestational ages and a possible role of extra-adrenal inhibitory factors have been proposed⁹⁻¹². The possible steroidogenic effect of α -melanocyte-stimulating hormone (α MSH) on fetal adrenocortical function is controversial¹³⁻¹⁶, and we could not demonstrate any such action in fetal lamb or rabbit^{17,18}. Prolactin and growth hormone (GH) potentiate the steroidogenic effect of ACTH in adult rats¹⁹⁻²² but prolactin had no such effect in fetal lamb¹⁷. We have investigated the steroidogenic properties of GH, and report here that it stimulates adrenal steroidogenesis in the fetal but not maternal rabbit *in vitro* and in fetal but not maternal sheep *in vivo*.

We used pregnant New Zealand White rabbits of known gestational age (g.a. 25-30 days, term 31 days, $n=8$) and Columbia-Suffolk strain ewes (g.a. 140-144 days, term 150 days, $n=5$). Each was killed by intravenous pentobarbitone and after hysterotomy, the fetuses were rapidly removed. Adrenal glands were removed *in toto*, without disrupting the capsule, and bathed in culture medium M199 with Hank's salt solution (Gibco) containing 2.2 mM calcium chloride. Adrenals were pooled within each litter of rabbits. Glands were trimmed of fat and connective tissue, minced when necessary and bathed in 5 ml of medium M199 containing 2.5 mg ml⁻¹ of lyophilized trypsin (Worthington), which digests adrenal medullary cells but leaves adrenocortical cells intact. We previously described¹⁸ the details of isolating dispersed viable adrenocortical cells by a modification of Sayers' technique²³. We then incubated 3×10^4 cells in 150 μ l (rabbit) or 10^5 cells in 200 μ l (sheep) of the culture

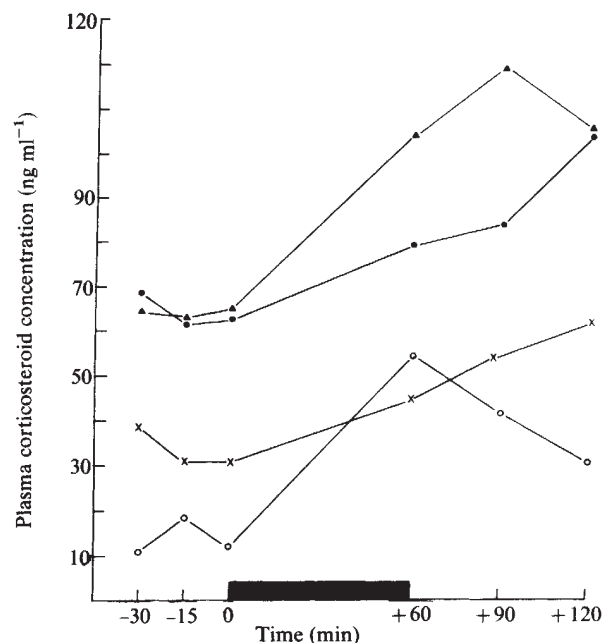


Fig. 1 Fetal lamb plasma corticosteroid concentrations before and after HGH infusion (shown by shaded bar). Age of lambs: ▲, 138 days; ●, 143 days; ×, 139 days; ○, 137 days.

medium, in duplicate with 50 μ l of phosphate buffer (0.01 M monobasic sodium phosphate, 0.9% NaCl, pH 7.4), adding human GH (HGH, National Pituitary Agency, NIH-HS 2243E) or synthetic ACTH₁₋₂₄ (Organon), the biologically active portion of the intact hormone ACTH₁₋₃₉, in the buffer. The incubation was performed in polystyrene tubes (12 × 75 mm) in a water bath at 37 °C under a mixture of 5% CO₂ and 95% O₂ for 2 h, after which the cells were removed by centrifugation at 300g for 10 min. Aliquots (50 μ l) of supernatant were removed, in duplicate, to measure corticoid production. Response to ACTH₁₋₂₄ served as the control for the steroidogenic response of the cell preparation. Corticosteroids were measured by a specific radioimmunoassay (RIA)^{24,25}. Because rabbit adrenals secrete both cortisol and corticosterone²⁶, an antibody (Radioassay System Catalogue 147) that reacts 100% with both hormones was used with rabbit adrenal cells, but because cortisol is the major corticoid in sheep²⁷, an antibody (Radioassay System Catalogue 137) that cross-reacts 100% with cortisol and minimally with other corticoids was used with sheep adrenal cells. The results represent total corticosteroid production by dispersed rabbit and sheep adrenocortical cells.

For *in vivo* studies, pregnant Columbia-Suffolk strain ewes of known gestational age and adult non-pregnant sheep were used. Pregnancy was confirmed by X-ray analysis between 85 and 90 days of gestation. The maternal and fetal carotid artery and jugular vein were catheterized²⁸ at least 1 week before the experiment. At 137, 138, 139 and 143 days of gestation, 20 μ g

Table 1 Basal corticosteroid production and increment above basal from hormone stimulation by fetal and maternal rabbit adrenocortical cells

Corticosteroid production (ng per 3×10^4 cells per 2 h)		
	Fetus ($n=8$)	Mother ($n=4$)
Basal	3.69 ± 0.66	3.38 ± 0.52
Increment above basal		
ACTH (25 μ g ml ⁻¹)	2.69 ± 0.32	2.80 ± 0.26
ACTH (125 μ g ml ⁻¹)	3.96 ± 0.47	3.85 ± 0.17
HGH (1.25 μ g ml ⁻¹)	3.30 ± 0.37	-0.05 ± 0.08
HGH (2.5 μ g ml ⁻¹)	4.59 ± 0.27	-0.04 ± 0.05

Values are mean $\pm$ s.e.m.

* To whom reprint requests should be addressed.

of HGH in 8 ml of normal saline was infused for 1 h through the fetal jugular vein. Fetal arterial blood (1 ml) was drawn at timed intervals before and after starting HGH infusion as indicated in Fig. 1; pH, pO₂ and pCO₂ were measured before and after infusion and were similar and in the normal range of our experimental animals. To exclude the effects of stress due to saline infusion, two animals earlier infused with HGH were infused with 8 ml of normal saline through the fetal jugular vein at 140 and 141 days of gestation and fetal corticosteroid concentration in plasma measured. We also infused HGH into catheterized adult sheep ($n = 3$) using the same protocol. With-drawn blood was immediately centrifuged at 500g for 10 min at 4°C and plasma stored at -20°C until measurement of corticosteroid concentrations²⁵. Plasma concentrations of HGH achieved *in vivo* were measured by RIA²⁹ in two fetal and one adult sheep before and after HGH infusion.

The mean basal production of corticoids, and the increase (mean $\pm$ s.e.m.) due to hormonal stimulation in rabbit and sheep adrenal cells are shown respectively in Tables 1 and 2. Although fetal cells from both species significantly increased steroidogenesis under the influence of ACTH₁₋₂₄ or HGH, the maternal cells responded to ACTH₁₋₂₄ only. Furthermore corticoid production by fetal rabbit and fetal sheep cells after incubation with 2.5 $\mu\text{g ml}^{-1}$ of HGH was significantly higher than after 1.25 $\mu\text{g ml}^{-1}$ (paired *t*-test $P < 0.02$).

Basal plasma corticosteroid concentrations in fetal lamb were significantly higher than in adult sheep (Figs 1, 2), but in both were within range reported by others³⁰⁻³². Plasma corticosteroid concentrations increased between 60 and 350% above basal after infusion of HGH in all four fetal lambs (Fig. 1). No such increase was noted in fetuses infused with saline or in adults infused with HGH (Fig. 2). Basal fetal and adult plasma HGH concentrations were less than the sensitivity of the assay (0.1 ng ml⁻¹), indicating minimal cross-reaction between endogenous ovine growth hormone and the HGH, especially as GH concentrations in ovine fetal plasma are high^{33,34}. Fetal lamb plasma HGH concentrations at the end of infusion (+60 min) were 15.3 and 25 ng ml⁻¹ while adult sheep plasma HGH concentration was 2.1 ng ml⁻¹.

The HGH we used is a highly purified immunochemical. Because maternal adrenal cells showed no steroidogenic response, unlike fetal cells, contamination of HGH with biologically active ACTH seemed unlikely. To confirm this assumption, the buffer and 2.5 μg of HGH in the buffer were assayed for ACTH by RIA at Nichols' Institute³⁵. The ACTH concentrations in HGH plus buffer were the same as those in the buffer alone. Contamination with αMSH seemed unlikely for the following reasons. Using the same dispersed cell system and RIA techniques, we have observed no steroidogenic effect of αMSH in concentrations from 8 to 800,000 pg ml⁻¹ in either fetal sheep or rabbits at various gestational ages^{17,18}. Similarly, infusion of 75 μg of synthetic αMSH (Bachem) in chronically catheterized fetal sheep at 127-130 days of gestation ($n = 4$) produced no increase above basal fetal plasma corticosteroid concentrations¹⁷. As a direct steroidogenic effect of HGH was observed *in vitro*, the increase in fetal plasma corticosteroid

Table 2 Basal corticosteroid production and increment above basal from hormone stimulation by fetal and maternal sheep adrenocortical cells

	Corticosteroid production (ng per 10 ⁵ cells per 2 h)	
	Fetus ($n = 5$)	Mother ($n = 3$)
Basal	1.79 $\pm$ 0.42	1.92 $\pm$ 0.59
Increment above basal		
ACTH (25 pg ml ⁻¹)	2.87 $\pm$ 0.63	5.04 $\pm$ 1.54
HGH (1.25 $\mu\text{g ml}^{-1}$)	3.08 $\pm$ 0.86	-0.15 $\pm$ 0.17
HGH (2.5 $\mu\text{g ml}^{-1}$)	5.33 $\pm$ 1.1	0.13 $\pm$ 0.19

Values are mean $\pm$ s.e.m.

observed *in vivo* is unlikely to have been due to ACTH released from the fetal pituitary.

Although no data are available for the rabbit, fetal plasma GH concentrations are very high compared with those during neonatal and adult life in sheep^{33,34} and humans³⁶. Because GH does not cross the placenta and the metabolic clearance rates in the fetal and neonatal period are similar³⁷, high fetal plasma concentrations probably represent increased secretion by the fetal adenohypophysis. The physiological significance of high fetal plasma GH concentrations is unclear; it does not seem to be essential for somatic growth³⁷. Our findings provide the first experimental evidence that GH directly influences adrenal steroidogenesis during fetal life. Although concentrations of HGH achieved *in vivo* were within the physiological concentrations of ovine GH observed in fetal (up to 250 ng ml⁻¹) and adult sheep plasma^{33,34}, the concentrations of HGH tested *in vitro* were one log order greater. Using entire adrenal glands rather than dispersed cells, collected from fetal rabbits at 24-27 days of gestation, Rudman *et al.* observed no increase in corticosterone production due to HGH stimulation¹⁶. The discrepancy with our results may be due to the different design of the experiments, the different gestation, the different RIA or a combination of these factors.

Received 9 June 1980; accepted 21 January 1981.

- Liggins, G. C. *Am. J. Obstet. Gynec.* **126**, 931 (1976).
- Thornburn, G. D. & Challis, J. R. G. *Physiol. Rev.* **59**, 863 (1979).
- Nathanielsz, P. W. *A. Rev. Physiol.* **40**, 411 (1978).
- Ballard, P. L. *Monographs on Endocrinology: Glucocorticoid Hormone Action* (eds Baxter, J. D. & Rousseau, G. G.) 493 (Springer, Berlin, 1979).
- Thomas, A. L., Crane, E. J. & Nathanielsz, P. W. *Endocrinology* **103**, 17 (1978).
- Sayers, G. & Portanova, R. *Handbook of Physiology* Vol. 5 Sect. F (eds Green, R. D. & Astwood, E. B.) 41 (Am. Physiol. Soc., Washington DC, 1975).
- Rees, L., Jack, P. M. B., Thomas, A. L. & Nathanielsz, P. W. *Nature* **253**, 274 (1975).
- Winters, A. J., Oliver, C., Colston, C., MacDonald, P. C. & Porter, J. C. *J. clin. Endocr. Metab.* **39**, 269 (1976).
- Glickman, J. A. & Challis, J. R. G. *Endocrinology* **106**, 1371 (1980).
- Madill, D. & Bassett, J. M. *J. Endocr.* **58**, 75 (1973).
- Roebuck, M. M., Jones, C. T., Holland, D. & Silman, R. *Nature* **284**, 616 (1980).
- Nathanielsz, P. W. *Fetal Endocrinology: An Experimental Approach*, 191 (North-Holland, Amsterdam, 1976).
- Lianos, A. J. *et al. Endocrinology* **105**, 613 (1979).
- Walsh, S. W., Norman, R. L. & Novy, M. J. *Endocrinology* **104**, 1805 (1979).
- Jujeda, K., Faiman, C., Reyes, F. I., Thliveris, J. & Winter, J. S. D. *Abstr. Endocrine Soc. Mtg* **626**, 229 (1979).
- Rudman, D., Hollins, B. M., Lewis, N. C. & Chawla, R. K. *J. clin. Invest.* **65**, 822 (1980).
- Magyar, D. M., Devaskar, U. P., Fridshal, D., Buster, J. E. & Nathanielsz, P. W. *Endocrinology* **107**, 1582 (1980).
- Devaskar, U. P., Magyar, D., Fridshal, D., Buster, J. E. & Nathanielsz, P. W. *Endocrinology* **107**, 809 (1980).
- Lis, M., Gilardeau, C. & Cheretien, M. *Clin. Res.* **21**, 1027 (1973).
- Colby, H. D., Caffrey, J. L. & Kitay, J. I. *Endocrinology* **93**, 188 (1973).
- Kramer, R. E., Crier, J. W. & Colby, M. D. *Endocrinology* **101**, 297 (1977).
- Coyne, M. D. *Abstr. Endocr. Soc. Mtg* **632**, 230 (1979).
- Sayers, G., Swallow, R. L. & Giordano, N. D. *Endocrinology* **88**, 1063 (1971).
- Radioassay Systems Laboratories, Inc. *Direct Corticosterone by RIA* (1979).
- Radioassay Systems Laboratories *Cortisol* ¹²⁵I Kit Cat. **137**, 133 (1978).
- Mulay, S., Giannopoulos, G. & Solomon, S. *Endocrinology* **93**, 1342 (1973).
- Thomas, S. J., Wilson, D. W., Pierrepont, C. G., Cameron, E. H. D. & Griffiths, K. T. J. *Endocr.* **68**, 181 (1976).
- Sperling, M. A., Christensen, R. A., Ganguli, S. & Anand, R. *Pediat. Res.* **14**, 203 (1980).
- Goodman, H. G., Grumbach, M. M. & Kaplan, S. L. *New Engl. J. Med.* **278**, 56 (1968).
- Strott, C. A., Sundel, H. & Stahlman, M. T. *Endocrinology* **95**, 1327 (1974).
- Brown, E. H., Coghlan, J. P., Hardy, K. J. & Wintour, E. M. *Acta endocr. Copenh.* **88**, 364 (1978).
- Magyar, D. M. *et al. Endocrinology* **107**, 155 (1980).
- Bassett, J. M., Thornburn, G. D. & Wallace, A. L. *C. J. Endocr.* **48**, 251 (1970).
- Gluckman, P. D., Mueller, P. L., Kaplan, S. L., Rudolph, A. M. & Grumbach, M. M. *Endocrinology* **104**, 162 (1979).
- Nichols, A. L. & Nelson, J. C. *Radioimmunoassay Manual* Vol. 1 (Nichols Institute, 1977).
- Kaplan, S. L., Grumbach, M. M. & Shephard, T. H. *J. clin. Invest.* **51**, 2080 (1972).
- Nathanielsz, P. W. *Fetal Endocrinology: An Experimental Approach*, 108 (North-Holland, Amsterdam, 1976).

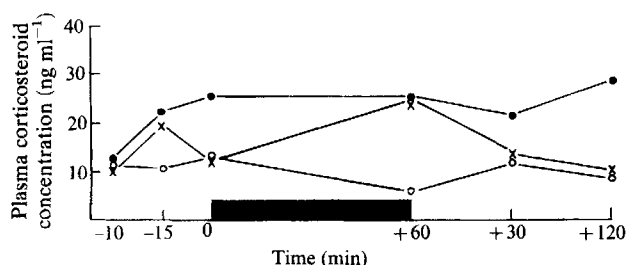


Fig. 2 Adult sheep plasma corticosteroid concentrations before and after HGH infusion (shown by shaded bar).

The respiratory burst of phagocytic cells is associated with a rise in vacuolar pH

Anthony W. Segal*, Michael Geisow†, Rodolfo Garcia*, Angela Harper* & Robert Miller‡

* Department of Haematology, School of Medicine, and

† Department of Biophysics, NIMR, Mill Hill, London NW7 1AA, UK

‡ Department of Zoology, University College London, London WC1E 6BT, UK

Neutrophil leukocytes are the body's major defence against bacteria, which they phagocytose and kill. It has been found that phagocytosis and killing are accompanied by a dramatic rise in non-mitochondrial respiration^{1,2}; and that the efficiency of killing is impaired in the absence of oxygen^{3,4}. It is also impaired in neutrophils from patients with chronic granulomatous disease (CGD), where the respiratory burst is absent⁵. This has been difficult to reconcile with their normal content of granule proteins that kill bacteria *in vitro*⁶. Indeed, CGD cells are essentially normal both morphologically and constitutionally^{7,8} except that they lack a functional very low potential cytochrome *b* (*b*₂₄₅)⁹, which is a component of the oxidase system responsible for the respiratory burst of normal cells^{10,11}. Activation of the oxidase is associated with the generation of various reduced oxygen species¹²⁻¹⁶ which have been widely thought to be responsible for the killing of phagocytosed microorganisms either directly¹⁷, or by acting as substrates for myeloperoxidase-mediated halogenation¹⁸. We report here, however, that a major consequence of the defective function of this oxidase in neutrophils and monocytes from CGD patients is an absence of the normal initial rise, and an unusually rapid and extensive fall in pH which is itself associated with the impairment of the killing and digestion of intracellular staphylococci.

Phagocytosis and bacterial killing were measured in cells from five severely affected male patients with X-linked CGD, all of whose cells lacked spectral evidence of cytochrome *b*₂₄₅ (ref. 9), a mildly affected girl with the autosomal recessive variant⁹ and normal subjects. Phagocytosis was assayed by lysostaphin lysis of extracellular organisms¹⁹. It was very rapid in normal cells (*t*_{1/2} = 0.65 min) and maximal at 4 min (Fig. 1). The bacteria then re-emerged from the cells as shown by an increase in extracellular radioactivity from 13% at 4 min to 32% at 16 min. This expulsion of phagocytosed material, mostly in the form of particles which sedimented after gentle centrifugation, has been described previously²⁰. Phagocytosis by the patients' cells was slightly slower (*t*_{1/2} = 1.26 min), and there was no evidence of the subsequent release of bacteria from the cells (Fig. 1).

Bacterial killing by normal cells was very rapid (Fig. 1). The patients' cells also killed more than half the organisms and when corrections were made for extracellular organisms it was calculated that no bacteria remained viable within normal cells and that only 22% of those within the patients' cells survived.

To what can one attribute the reduced efficiency of bacterial killing by CGD cells? We have previously proposed that one role of the vacuolar cytochrome system is to regulate the pH within phagocytic vacuoles^{11,21}, and there is evidence that the pH falls faster after phagocytosis in CGD than in normal cells²². Thus, we measured early pH changes in the phagocytic vacuoles of leukocytes from healthy and granulomatous individuals, using the pH indicator fluorescein²³ conjugated to *Staphylococcus aureus*. Glass coverslips were coated with monolayers (1 × 10⁴ cells mm⁻²) of neutrophils (>95% pure) and mononuclear cells (>80% monocytes) which were then exposed to a suspension of fluoresceinated bacteria at 4 °C. The pH of the bacterial environment was determined by measuring fluorescein emission

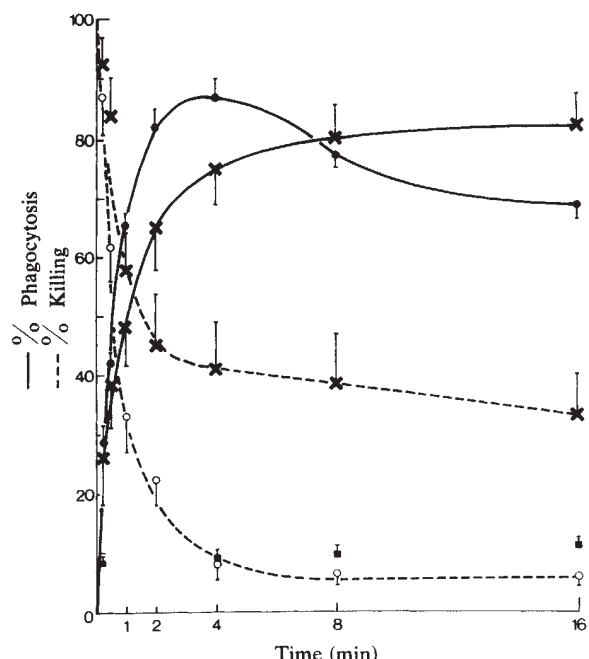


Fig. 1 Phagocytosis and killing of staphylococci by normal (●, ○) and CGD (×) neutrophils. Solid line, intracellular bacteria; broken line, viable bacteria (%). *S. aureus* (Oxford Strain NCTC 6571) from an overnight culture in broth (Oxoid) containing 25 μCi ml⁻¹ ³H-glucose (Radiochemical Centre, TRK. 85) were opsonized with human IgG (ref. 41) and added (in a 1:1 ratio) to 1 × 10⁸ purified neutrophils (>95% in normals, patients' cells containing 5–10% eosinophils) in 1.5 ml RPMI 1640 medium (Flow Labs) containing 5 IU ml⁻¹ heparin in a rapidly stirred thermostatically controlled chamber⁴¹. Aliquots (50 μl) were taken into 500 μl ice-cold RPMI containing *N*-ethyl maleimide (1 mM) and 100 μg ml⁻¹ lysostaphin (Sigma). The samples were mixed, incubated at 37 °C for 15 min, re-cooled and centrifuged at 7,000g for 4 min. Radioactivity in the supernatant was measured in a liquid scintillation counter. Fifty-microlitre aliquots of the mixture of bacteria and cells were taken at the same times into 10 ml ice cold H₂O. After 1 h aliquots were plated out for colony counts. Results are expressed as the mean ± s.e. of seven studies on cells from six CGD patients, each study being controlled with a simultaneous study on normal cells. Lactate dehydrogenase (mean ± s.e. of four experiments) release³⁰ from normal cells is shown (■).

at selected excitation wavelengths at various times after the initiation of phagocytosis. In normal neutrophils an increase in peak fluorescence corresponding to an apparent rise of intravacuolar pH to 7.75 was observed within the first 2 min (Fig. 2). Thereafter the pH fell slowly to 6.0–6.5 after 2 h. The same pattern was observed with normal monocytes (Fig. 2b). Similar studies with yeasts impregnated with the pH indicator bromothymol blue^{22,24} also showed that almost immediately following phagocytosis by normal cells there was a transient colour change from yellow to blue indicating a rise in pH from 6.8 in the medium to >7.6.

The pattern of pH changes with phagocytosis was clearly different in CGD leukocytes. There was neither colour change with bromothymol blue-impregnated yeasts, nor initial fluorescence increase with fluorescein-labelled *S. aureus*; instead the pH fell rapidly from 7.4 to 6.7 within 2 min. The final vacuolar pH of 5.5 was lower than that of normal cells. Addition of methylamine, which elevates the pH of secondary lysosomes²³, to the cell perfusate and the pre-incubation medium, effectively prevented this early rapid drop in vacuolar pH (Fig. 2), although phagocytosis was not inhibited.

Furthermore, spectra from fluorescent *S. aureus* bound to CGD or normal neutrophils at 10 °C (before phagocytosis) were identical to the spectrum of the free bacterium at pH 7.4, whereas after 13 min incubation at 37 °C the spectra were

completely different; that of CGD cells could be fitted best by the pH 5.9 spectrum of fluorescein and that of normal cells by the pH 7.0 spectrum (data not shown).

When the experiment was repeated with normal human neutrophils in anaerobic conditions, achieved by perfusing the cells with degassed buffers or mixing a concentrated suspension of neutrophils with the fluorescent bacteria at 4 °C in fine capil-

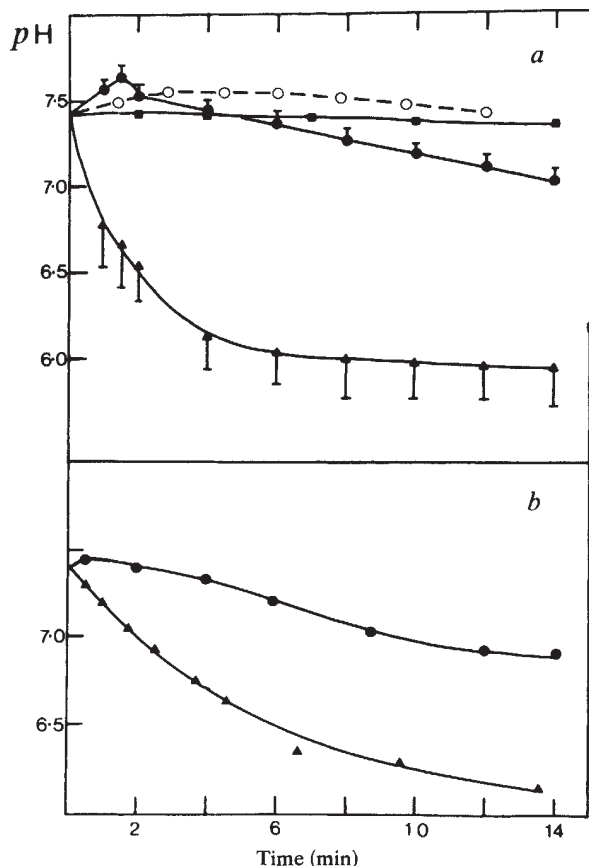


Fig. 2 Phagosomal pH in leukocytes after phagocytosis. *a*, Neutrophils: mean $\pm$ s.d. of two studies on each of two patients with CGD and one study on a third ($\blacktriangle$) and on five paired control subjects ($\bullet$). The effects of methylamine²³ (10 mM) on the pH within CGD cells ($\blacksquare$) and cytochalasin β ($1 \mu\text{g ml}^{-1}$) on control cells ($\circ$) are also shown. *b*, Monocytes: mean results of studies on two CGD patients ($\blacktriangle$) and paired control subjects ($\bullet$). *S. aureus* (50 μl of packed organisms) was well washed with saline and 50 mM pyridine-HCl buffer, pH 5.0. The bacteria were resuspended in 1 ml of the same buffer and mixed with 40 μl of fluorescein amine isomer I (Sigma), 1 mg ml^{-1} in dioxan. One mg of 1-ethyl 3-diaminopropyl carbodiimide was added and the suspension agitated for 2 h. The suspension was washed first in the pyridine buffer containing 100 mM glycine and then in saline. Fluorescence microscopy showed that the killed bacteria had been uniformly surface labelled. Freshly isolated human neutrophils or mononuclear cells⁴² were allowed to attach to acid-etched glass coverslips for 15 min at room temperature. Non-adherent cells were then washed away with RPMI medium, and the resulting monolayers were stored under this medium. Fluorescent *S. aureus* opsonized with IgG (ref. 41) were allowed to sediment on to leukocyte monolayers and attach at 4 °C for 20 min. Free or loosely bound labelled bacteria were removed by vigorous washing in ice-cold saline or RPMI medium. Fluorescence monitoring with a Perkin-Elmer MPF 4 fluorimeter began immediately after transfer of the cold monolayer to pre-warmed fluorescence cuvettes perfused with oxygenated RPMI medium at 37 °C. Phagosomal pH was calculated from the shape of the pH-dependent excitation spectrum of fluorescein by comparison with calibration spectra over a pH range of 4–8 from the same fluorescent bacteria adsorbed to glass in the absence of cells. The excitation spectra from fluorescent bacteria were corrected for the intrinsic fluorescence of the leukocytes and for scattered light.

laries which were then sealed, the initial pH rise was abolished and the pH fell rapidly, just as in the CGD patients' cells.

In these experiments bacteria were layered on the monolayers of cells at 4 °C to permit attachment without phagocytosis, which was then synchronized by rapid warming to 37 °C. To ensure that cooling the cells to 4 °C was not modifying the results, the cells were incubated with bacteria at 25 °C for 30 s before vigorous washing and pH measurement. The same pH changes as described above were observed in normal and CGD cells. Pre-incubation with cytochalasin B (5 μM) retarded phagocytosis and diminished and delayed the pH changes (Fig. 2). Neither the direct action of the products of the oxidase, generated at pH 7.4 with a xanthine oxidase-hypoxanthine system⁹, nor the method of coupling the fluorescein to the organism, on carboxyl groups with carbodiimide or on amino groups with isothiocyanate, altered the fluorescence characteristics of fluorescein.

The fluorescence measurements include signals from both phagocytosed bacteria and extracellular organisms which remain in the buffered medium. As the constant signal from the latter minimizes the average change, and consequently that attributable to intracellular particles, it was important to measure the phagocytosis of bacteria by the monolayer cells, which was done by the release of radioactivity from extracellular bacteria after digestion with lysostaphin as described in Fig. 1 legend. Fifty per cent were taken up by 30 s, 80% by 2 min and a maximum of 10% finally remained extracellular. The kinetics of phagocytosis were the same in cells from a patient with CGD. In studies with fluorescent *S. aureus* the addition of lysostaphin after 15 min caused a small decrease in fluorescence, but insufficient to alter the pH profiles, confirming that the pH measurement pertained to organisms within the cell.

To determine whether the excessively acid conditions within the vacuole of CGD cells could explain the defective bacterial killing, we examined the killing of staphylococci by the cells of two patients in the presence and absence of 10 mM methylamine, which we had shown to elevate the pH in the vacuoles of patients' cells (Fig. 2). Regression analysis of the relationship between extracellular and viable organisms (see Fig. 1 legend) indicated that in the presence of methylamine (which did not itself kill bacteria) all the intracellular organisms were killed and that bacterial killing was significantly better ($P < 0.02$)²⁵ than in its absence.

To investigate other possible effects of the excessive acidity in CGD vacuoles we measured the digestion of radiolabelled, heat-killed staphylococci. Solubilization of the phenylalanine label was less than half of normal, that of the incorporated glucose was about 80% of normal whereas that of uridine label was normal (Fig. 3). Phagocytosis of the labelled bacteria by the two groups of cells was similar. The differences in the solubilization of bacteria seemed unrelated to differences in their content of hydrolytic enzymes because incubation of labelled bacteria with homogenates of normal or CGD cells at pHs of 5–8 released similar amounts of radiolabelled phenylalanine and glucose. Nor could the reduced solubilization in CGD cells be accounted for by less efficient degranulation. Phagocytic vacuoles, isolated by flotation²⁶, from four patients were in fact found to contain more enzymes than normal, the specific activities of β glucuronidase²⁷, lysozyme²⁸ and myeloperoxidase²⁹ being 4.55, 2.86 and 1.19 of those in vacuoles from the paired control cells. Half-maximal degranulation times for these enzymes, determined over 1 h, were normal²⁶.

We next investigated the difference between the expulsion of phagocytosed bacteria by normal and CGD cells (Fig. 1). We found that normal cells released three times more solubilized ³H-glucose label than cells of the four patients (Fig. 3a, $P < 0.0025$)²⁵, although cell damage as measured by lactate dehydrogenase release³⁰ after incubation for 2 h was not significantly different (5.7 ± 1.2 (s.e.) in controls and $8.0 \pm 3.9\%$ in patients). For morphological studies cells and bacteria were incubated in the conditions used in the killing assay (see Fig. 1 legend) and timed samples were collected for electron

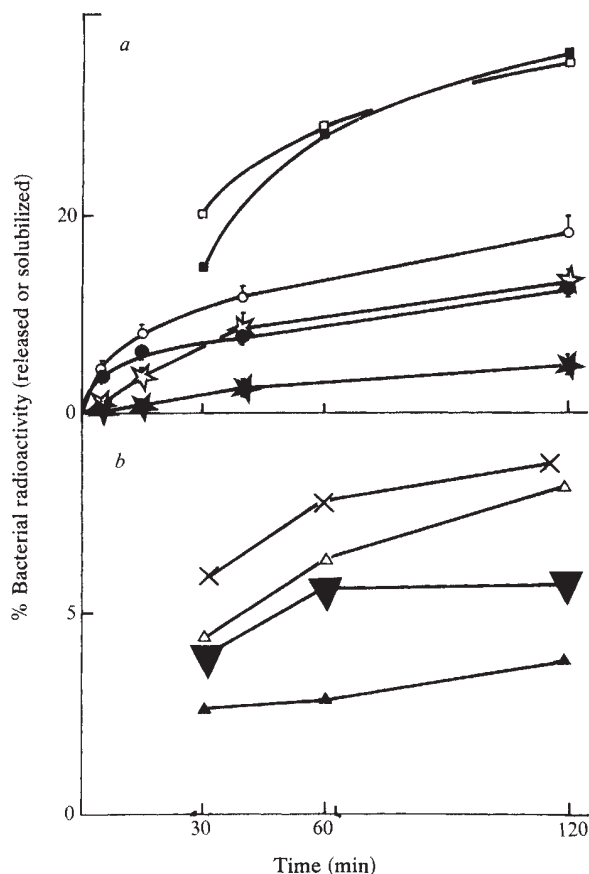


Fig. 3 Solubilization of heat-killed radiolabelled staphylococci by, and release of solubilized radioactivity from, CGD and normal neutrophils. Bacteria were grown in broth containing ^3H -glucose ($25 \mu\text{Ci ml}^{-1}$), ^3H -uridine ($1 \mu\text{Ci ml}^{-1}$) or ^{14}C -phenylalanine ($0.2 \mu\text{Ci ml}^{-1}$) (Radiochemical Centre) and then washed, killed by heating in a boiling water bath for 10 min, rewashed and opsonized with IgG⁴¹. Radiolabelled bacteria were mixed with cells (1:1) and rapidly shaken at 37°C . Solubilization was measured by lysing the cells with distilled water, centrifuging at $8,000g$ for 10 min, homogenizing the pellet in distilled water recentrifuging and counting the radioactivity in the supernatants. Release of solubilized radioactivity was examined by centrifuging the cells and bacteria at $7,000g$ for 1 min after various incubation times, and determining the radioactivity in the supernatant. *a*, Solubilization of uridine (mean of two patients, ■, and controls, □) and glucose (mean $\pm$ s.e. of four patients, ●, and controls, ○) and release of the solubilized ^3H -glucose label (mean $\pm$ s.e. of four patients, ★, and controls, ☆). *b*, Solubilization of ^{14}C -phenylalanine by cells from two patients (▲) and controls (Δ). Methylamine (10 mM) enhanced solubilization by both the CGD (▼) and control (×) cells.

microscopic examination. After about 2 min the vacuoles in normal, but not in CGD, cells began to swell and became much larger than the enclosed bacterium (Fig. 4). The mean vacuolar volume in normal cells, as calculated from the cross-sectional area of vacuoles on electron micrographs, was approximately $15 \mu\text{m}^3$ after 8 min and $22 \mu\text{m}^3$ after 16 min, in comparison with 4.0 and $4.5 \mu\text{m}^3$ in the patients' cells after the same incubation times (see Fig. 4 legend).

The abnormally small size of phagocytic vacuoles in CGD^{8,31} has been attributed to a reduction in the release of granule contents into the vacuole as a result of defective degranulation. However, as described here and by others³², degranulation is not defective in CGD and although the enlarged vacuoles seen in normal cells appear swollen, they are not obviously distended with electron-dense granule contents. We propose instead that phagocytic vacuoles in normal cells enlarge as a result of the attraction of water by osmotically active digestion products of

the ingested bacteria, and that the vacuoles in CGD cells are much smaller because, as shown in Fig. 3, digestion within their vacuoles is impaired. This hypothesis is supported by our findings that even after incubation times as long as 1 h there is virtually no swelling of vacuoles by normal cells after the phagocytosis of indigestible latex particles opsonized with the same IgG as the bacteria, and that in the presence of methylamine, which enhances bacterial digestion, the mean volume of the patients' vacuoles increased from a calculated 4.5 to $15 \mu\text{m}^3$ ($P < 0.005$, see Fig. 4 legend).

We suggest that the normal sequence of events following phagocytosis of staphylococci is as follows. After entry of the bacteria, the vacuole oxidase system elevates the intravacuolar pH via a cytochrome b_{-245} mediated⁹⁻¹¹ transport of electrons from the cytosolic to the luminal surface. The resultant reduction of molecular oxygen to O_2^{2-} would consume protons during the formation of H_2O_2 (refs 13, 14), raising the intravacuolar pH. The raised pH facilitates killing and bacteriolysis by granule proteins³³⁻³⁶ following which the pH of the vacuoles is reduced to optimize the activities of hydrolases and other proteins with acid pH optima³⁷. In normal vacuoles the generation of low-molecular-weight hydrolysis products is evidently sufficiently rapid to swell osmotically the vacuoles containing the disintegrating organisms. The vacuole enlarges until it comes into contact and fuses with the plasma membrane, releasing the partially digested bacterium and solubilized products, possibly with immunogenic potential, into the extracellular environment. In CGD, however, the absence of the oxidase results in a rapid fall in vacuolar pH inhibiting the killing mechanisms, although not before they have disposed of most of the ingested organisms, and impairing subsequent microbial digestion and its effects on vacuolar morphology. Impaired digestion could account for the lipochromic material that accumulates within histiocytes in the tissues of these patients^{38,39}.

Differences in the control of the pH of endocytic vacuoles might be one of the factors responsible for variations in the susceptibility of individuals to infection and could be manipulated therapeutically⁴⁰.

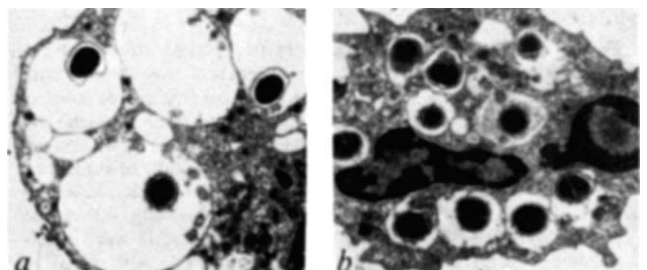


Fig. 4 Transmission electron micrographs of neutrophils from *a*, a normal subject, and *b*, a male patient with CGD ($\times 5,600$). Samples were fixed and processed 16 min after exposure to opsonized staphylococci. Phagocytic vacuoles in normal cells were found to be enlarged and to contain bacterial debris whereas those in CGD cells were much smaller and the bacteria appeared undegraded. Prints were made of electron micrographs of 10 randomly selected cells at each time interval and between 20 and 40 phagocytic vacuoles containing bacteria that appeared to have been sectioned through their region of maximum diameter were excised and weighed. The samples were coded and only identified after the conclusion of the measurements. Cross-sectional areas were determined by comparing the weights of excised vacuoles with those of intact bacteria (radius $\sim 0.6 \mu\text{m}$)⁴³. The cross-sectional areas of the patients' vacuoles were significantly smaller than those of the control (2.8 ± 0.2 and 2.9 ± 0.2 compared with 6.5 ± 1.2 and 8.3 ± 1.0 times that of the bacteria at 8 and 16 min respectively, $P < 0.005$ and $P < 0.001$)²⁵. The cells of another patient were examined and although not quantified, the vacuoles were also obviously much smaller than those in the paired control cells.

We thank the Wellcome Trust and MRC for financial support, Drs David Fellingham and Richard Whitton for helpful advice, and Professors D. Weatherall and M. Greaves for allowing us to study patients under their care.

Received 4 September 1980; accepted 10 February 1981.

1. Baldrige, C. W. & Gerard, R. W. *Am. J. Physiol.* **103**, 235–236 (1933).
2. Sbarra, A. J. & Karnovsky, M. L. *J. biol. Chem.* **234**, 1355–1362 (1959).
3. Selvaraj, R. J. & Sbarra, A. J. *Nature* **211**, 1272–1276 (1966).
4. Mandell, G. L. *Infect. Immun.* **9**, 337–341 (1974).
5. Holmes, B., Page, A. R. & Good, R. A. *J. clin. Invest.* **46**, 1422–1432 (1967).
6. Quie, P. G., White, J. G., Holmes, B. & Good, R. A. *J. clin. Invest.* **46**, 668–679 (1967).
7. Segal, A. W. & Peters, T. J. *Q. J. Med.* **47**, 213–220 (1978).
8. Klebanoff, S. J. & Clark, R. A. in *The Neutrophil*, 641–709 (North-Holland, Amsterdam, 1978).
9. Segal, A. W. & Jones, O. T. G. *FEBS Lett.* **110**, 111–114 (1980).
10. Segal, A. W. & Jones, O. T. G. *Nature* **276**, 515–517 (1978).
11. Segal, A. W. & Jones, O. T. G. *Biochem. J.* **180**, 33–44 (1979).
12. Babior, B. M., Kipnes, R. S. & Curnutte, J. T. *J. clin. Invest.* **52**, 741–744 (1973).
13. Iyer, G. Y. N., Islam, M. F. & Quastel, J. H. *Nature* **192**, 535–541 (1961).
14. Root, R. K., Metcalf, J., Oshino, N. & Chance, B. *J. clin. Invest.* **55**, 945–955 (1975).
15. Tauber, A. I. & Babior, B. M. *J. clin. Invest.* **60**, 374–379 (1977).
16. Green, M. R., Hill, H. A. O., Okolow-Zubrowska, M. J. & Segal, A. W. *FEBS Lett.* **100**, 23–26 (1979).
17. Babior, B. M., Curnutte, J. T. & Kipnes, R. S. *J. Lab. clin. Med.* **85**, 235–244 (1975).
18. Klebanoff, S. J. *Semin. Hemat.* **12**, 117–142 (1975).

19. Tan, J. S., Watanakunakorn, C. & Phair, J. P. *J. Lab. clin. Med.* **78**, 316–322 (1971).
20. Berlin, R. D., Fera, J. P. & Pfeiffer, J. R. *J. clin. Invest.* **63**, 1137–1144 (1979).
21. Segal, A. W. & Allison, A. C. *Ciba Fdn Symp.* **65**, 205–223 (1979).
22. Jacques, Y. V. & Bainton, D. F. *Lab. Invest.* **39**, 179–185 (1978).
23. Ohkuma, S. & Poole, B. *Proc. natn. Acad. Sci. U.S.A.* **75**, 3327–3331 (1978).
24. Mandell, G. L. *Proc. Soc. exp. Biol. Med.* **134**, 447–449 (1970).
25. Armitage, P. in *Statistical Methods in Medical Research*, 281–282, 289–291 (Blackwell, Oxford, 1971).
26. Segal, A. W., Dorling, J. & Coade, S. *J. Cell Biol.* **85**, 42–59 (1980).
27. Barrett, A. J. & Heath, M. F. in *The Lysosomes. A Laboratory Handbook* (ed. Dingle, T. T.) 19–145 (North-Holland, Amsterdam, 1977).
28. Gorin, G., Wang, S. F. & Papapavlou, L. *Ann. Biochem.* **99**, 113–127 (1971).
29. Bretz, U. & Baggolini, J. *Cell Biol.* **63**, 251–269 (1974).
30. Davies, P., Page, R. C. & Allison, A. C. *J. exp. Med.* **139**, 1262–1282 (1974).
31. Holmes, B., Quie, P. G., Windhorst, D. B. & Good, R. A. *Lancet* **i**, 1225–1228 (1966).
32. Stossel, T. P., Root, R. K. & Vaughan, M. *New Engl. J. Med.* **286**, 120–123 (1972).
33. Skarnes, R. C. & Watson, D. W. *J. exp. Med.* **104**, 829–845 (1956).
34. Skarnes, R. C. & Watson, D. W. *Bact. Rev.* **21**, 273–294 (1957).
35. Skarnes, R. C. *Nature* **216**, 806–808 (1967).
36. Odeberg, H. & Olsson, I. *J. clin. Invest.* **56**, 1118–1124 (1975).
37. Hirsch, J. G. *J. exp. Med.* **103**, 613–621 (1956).
38. Landing, B. H. & Shirkey, H. S. *Paediatrics* **20**, 431–438 (1957).
39. Clawson, C. C., Rodey, G. E. & Good, R. A. *Lab. Invest.* **22**, 294–300 (1970).
40. de Duve, C. *et al. Biochem. Pharmac.* **23**, 2495–2531 (1974).
41. Segal, A. W. & Coade, S. B. *Biochem. biophys. Res. Commun.* **84**, 611–617 (1978).
42. Boyum, A. *Scand. J. clin. Lab. Invest.* **21**, Suppl. 97, 31–50 (1968).
43. Luria, S. E. in *The Bacteria* (eds Gunsalus, I. C. & Stamer, R. Y.) **1**, 8 (Academic, New York, 1960).

Deficiency in cells expressing terminal transferase in autoimmune (motheaten) mice

Kenneth S. Landreth*, Kathleen McCoy†, James Clagett‡, F. J. Bollum§ & Cornelius Rosse*

Departments of *Biological Structure, †Microbiology/Immunology, and ‡Periodontics and Microbiology/Immunology, University of Washington School of Medicine, Seattle, Washington 98195, USA
§Department of Biochemistry, Uniformed Services University of the Health Sciences, Bethesda, Maryland 20014, USA

The extensive breakdown of immune homeostasis in the motheaten mouse (*me/me*) has been ascribed to a single gene defect on chromosome 6 (ref. 1). These mice develop skin lesions within the first week of life, do not thrive, and die within the first 3–8 weeks². There is severe hypergammaglobulinaemia³ with multiple species of circulating autoantibody^{3–5} and deposition of immune complexes in the thymus, skin, lungs and kidneys^{1,3}. A single gene defect producing such catastrophic results may provide an important model for understanding autoimmune phenomena. We report here a virtual absence of terminal deoxynucleotidyl transferase-positive (TdT⁺) cells in the bone marrow, thymus and spleen of motheaten mice. TdT is a DNA polymerase which has the unique capacity to polymerize nucleotides in the absence of template direction⁶. Although no *in vivo* biological function of this enzyme has been established, its unique appearance in the bone marrow and thymus of adult mammals and its *in vitro* biochemical activity⁶ have led to a proposed role for TdT in the somatic diversification of lymphocytes⁷. Bone marrow TdT⁺ cells have been shown to belong to both T⁸ and B^{9,10} cell populations and may also include precursor cells common to these lineages¹¹. Although the role of TdT in the acquisition of appropriate T- and B-cell specificities is not known, our results are the first to correlate the virtual absence of TdT⁺ cells with a severe autoimmune syndrome. We investigated the level of TdT⁺ cells in neonatal *me/me* mice and their normal littermates and the susceptibility of TdT⁺ cells to circulating autoantibody in motheaten mouse serum.

Bone marrow, thymus and spleen cells obtained from 21–25-day-old C3HeB/FeJ *me/me* mice, their normal littermates and C3HeB/FeJ wild-type mice were stained for TdT with immunoadsorbent-purified rabbit anti-calf TdT, previously shown to have good cross-reactivity with mouse TdT¹². The

TdT-specific reagent was visualized with fluorescein isothiocyanate (FITC)-conjugated F(ab')₂ goat anti-rabbit IgG. We found an extremely low incidence of TdT⁺ cells in both bone marrow and thymus of *me/me* mice compared with normal littermates and wild-type mice (Table 1). Most of the motheaten mice examined had no TdT⁺ cells, but when present, the numbers of these cells seemed to be inversely correlated with the severity of disease as judged by body weight and skin lesions (unpublished data).

As *me/me* mice possess circulating IgM autoantibodies to thymocytes^{1,5,13}, we tested whether or not motheaten mouse serum had autoantibody to TdT⁺ cells in the lymphoid tissues. Serum was collected from 10 *me/me* mice and 10 normal littermates between 21–25-days old, and incubated with C3HeB/FeJ wild-type bone marrow and thymus cells (Table 2). The binding of *me/me* serum IgM to normal bone marrow and thymus cells was visualized with tetramethylrhodamine isothiocyanate (TRITC)-conjugated specific goat anti-mouse μ (ref. 14) and TdT was stained as already described. As shown in

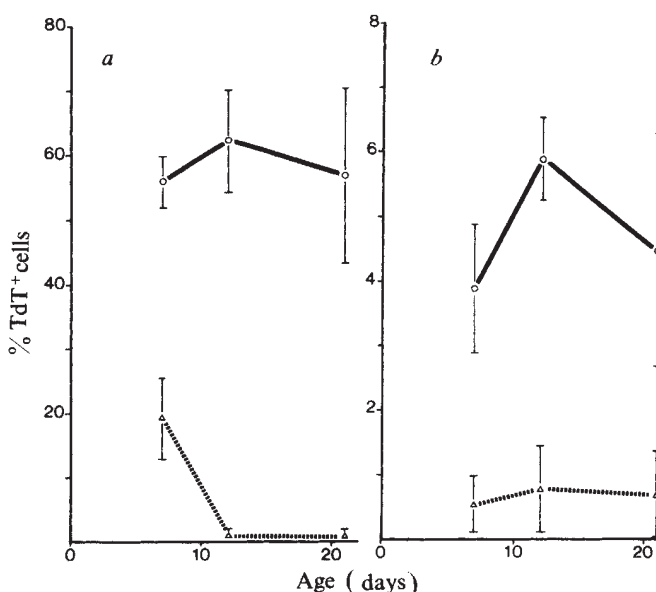


Fig. 1 The incidence of TdT⁺ cells in the thymus (a) and bone marrow (b) of *me/me* mice (Δ) and their normal littermates (○). Cells were collected and stained as described in Table 1. Over 1,000 cells were counted in cytocentrifuge preparations from each of five individual mice; values plotted are mean ± s.d.

† Present address: Sloan-Kettering Institute for Cancer Research, 145 Boston Post Road, Rye, New York 10580, USA.

Table 2, motheaten mice possess circulating IgM autoantibodies that bind to $64.28 \pm 10.37\%$ of normal thymocytes and to $15.38 \pm 7.54\%$ of bone marrow cells; IgM in normal littermate serum samples did not bind to any cells. In the thymus, most of the cells that bound *me/me* autoantibody were also TdT⁺, in contrast to the bone marrow, where none of the cells that bound autoantibody were TdT⁺.

The virtual absence of TdT⁺ cells in the thymus of *me/me* mice in the terminal stages of the autoimmune disease may be due to cytolysis by circulating autoantibody. In the thymus of normal littermates, $56.2 \pm 2.5\%$ of cells were TdT⁺ by 7 days of age and the incidence of these cells remained almost constant up to day 21 (Fig. 1; compare with ref. 6, Fig. 1), but in *me/me* mice only $19.6 \pm 6.8\%$ of thymus cells were TdT⁺ at 7 days. By day 12 the incidence of TdT⁺ cells in the thymus of *me/me* mice had fallen to $0.45 \pm 0.19\%$ and remained at this barely detectable level up to day 21. In normal littermates $3.89 \pm 1.01\%$ of bone marrow cells were TdT⁺ at day 7, increased to $5.91 \pm 0.67\%$ by day 12 and remained at $4.42 \pm 2.10\%$ on day 21. In *me/me* mouse bone marrow only 0.54 ± 0.45 – $0.76 \pm 0.75\%$ of nucleated cells were TdT⁺ at all time points of observation.

One explanation for the virtual absence of TdT⁺ cells in the thymus of motheaten mice by the age of 3 weeks could be autoimmunolysis of TdT⁺ thymocytes—such cell death would be consistent with the decrease in total thymic cellularity during the first 3 weeks¹. Our data suggest that the primary targets for the autoantibody are the TdT⁺ thymocytes found in the thymus cortex¹². Unlike the thymus, the number of TdT⁺ cells in motheaten marrow was low throughout the observation period. Because no binding of autoantibody to TdT⁺ cells could be demonstrated in the bone marrow of 21–25-day-old wild-type mice, it is difficult to attribute the absence of TdT⁺ cells in the marrow to autoimmunolysis. There are four possible explanations for this: (1) cells in defective bone marrow produce a variant of TdT not recognized by our antibody; (2) TdT⁺ cells

Table 2 Percentage of bone marrow and thymus cells of wild-type C3HeB/FeJ mice that bind autoantibodies

	Normal littermate serum ⁺	<i>me/me</i> serum ⁺	<i>me/me</i> serum ⁺ and TdT ⁺
Bone marrow	0	15.38 ± 7.54	0
Thymus	0	64.28 ± 10.37	52.50 ± 17.3

To assay circulating autoantibodies in *me/me* mice, bone marrow and thymus cells from wild-type C3HeB/FeJ mice were first incubated with intact or F(ab')₂ goat anti-mouse μ specific antibody for 30 min on ice to block μ determinants on the cell surface. Cells were then washed three times in HBSS (5% FCS, 0.08% NaN₃). Blocking efficiency was checked by incubation of these cell samples with TRITC-conjugated F(ab')₂ goat anti-mouse μ ¹⁴ for 30 min on ice followed by two washes of the cells with HBSS. This procedure routinely reduced detection of μ ⁺ cells to <0.5% in the marrow. One aliquot of the cell samples was mixed with normal mouse serum (diluted 1:5 in HBSS) and another aliquot with *me/me* mouse serum (diluted 1:5 in HBSS) and both were incubated for 30 min on ice. Both cell aliquots were then washed twice in HBSS followed by incubation with TRITC-F(ab')₂ goat anti-mouse μ as described. Washed cell aliquots were fixed in paraformaldehyde as described in Table 1 legend, cytocentrifuged on to clean glass slides and stained for TdT (see Table 1). Cell preparations were mounted in Tris-buffered glycerol, pH 7.0, and counted on a Leitz Ortholux epifluorescent microscope with appropriate filtration for scoring of both FITC-labelled TdT and TRITC-labelled IgM autoantibodies on the same slide. Cell counts represent >1,000 cells counted for each of 10 preparations; mean $\pm$ s.d.

are normally thymus derived; (3) TdT⁺ cells are rapidly exported from motheaten marrow to the involuting thymus or (4) there is a defective production of TdT⁺ bone marrow cells in these mutant mice. The first of these explanations seems unlikely because it requires postulating a species of TdT in *me/me* bone marrow that is antigenically distinct from TdT in *me/me* thymus or from that in normal mouse bone marrow and thymus cells. The post-thymic origin of TdT⁺ bone marrow cells is also considered unlikely because such cells are present in normal numbers in nude mouse bone marrow¹⁵. Present data do not allow a choice between the rapid export and deficient production of TdT⁺ cells. However, even if the paucity of TdT⁺ cells in the bone marrow is due to accelerated migration to the thymus, this explanation would not account for the absence of those TdT⁺ cells in the marrow which are committed to the B-cell lineage, because such cells are unlikely to migrate to the thymus.

The gross abnormality in immunoglobulin production that characterizes the *me/me* syndrome may well be linked to the defect in TdT⁺ B-lineage cells, whereas other manifestations of the *me/me* syndrome may depend on defects in TdT⁺ T cells of both bone marrow and thymic origin. Defective production of TdT or of TdT⁺ cells in the *me/me* mouse may provide important clues about autoimmunity and the normal biological function of TdT. If the enzyme does play a part in the somatic diversification of the molecules by which lymphocytes recognize antigen, then a failure in its production in the developing cell lineage might result in severely decreased heterogeneity of those molecules and in the antigen recognition capacity of the *me/me* mouse. Twenty-one-day-old *me/me* mice fail to respond to foreign antigenic stimulation by either specific antibody production^{3,13,16} or the generation of cytotoxic killer cells¹⁶, yet they exhibit hypergammaglobulinaemia^{1,3,16} and normal numbers of peripheral T cells^{1,5,16}. Preliminary data from our laboratory show that serum immunoglobulins from *me/me* mice have a lower heterogeneity than immunoglobulin from normal littermate serum analysed on two-dimensional polyacrylamide gels after immunoprecipitation of radiolabelled immunoglobulins (D. Putnam, K. McCoy, A. Jones and J. Clagett, unpublished data). These results strengthen the hypothesis that the underlying cause of motheaten pathology may, in part, result from a defect in the production of TdT.

Table 1 Incidence of TdT-positive cells in the thymus, bone marrow and spleen of motheaten and normal mice

	% TdT ⁺ cells		
	Thymus	Bone marrow	Spleen
Wild type	52.50 ± 12.34	3.66 ± 1.49	0.12 ± 0.27
Normal littermates	54.25 ± 19.99	4.21 ± 2.08	4.66 ± 2.13
Motheaten	1.43 ± 1.89	0.72 ± 0.81	0.61 ± 0.54

Motheaten mice on a C3HeB/FeJ background, normal littermates and wild-type C3HeB/FeJ mice (obtained from our own breeding colony or from Jackson Laboratories) were killed by cervical dislocation, the femora removed and the contents flushed at 4°C with Hanks' balanced salt solution (HBSS, Microbiological Associates) which contained 5% heat-inactivated fetal calf serum (FCS, Microbiological Associates) and 0.08% NaN₃. Spleens were placed in HBSS at room temperature and teased to extrude cells. Thymic lobes were minced in cold HBSS. Single cell suspensions of each tissue (>90% viable as assessed by trypan blue exclusion) were washed twice in HBSS. An equal volume of 2% paraformaldehyde in phosphate-buffered saline (PBS, pH 7.2) was added to 1.5 – 2.0×10^6 bone marrow, spleen and thymus cells in 0.25 ml HBSS. The cell suspension was kept on ice for 15 min followed by cytocentrifugation (Shandon Scientific) on to clean glass slides at 500 r.p.m. for 3 min. Slide preparations were air dried and immediately fixed in absolute methanol for 20 min at 4°C, then rehydrated through three changes of PBS containing 0.05% NaN₃ at 4°C and dried, excluding the area of the fixed cells which remained wet throughout the staining procedure. Monospecific rabbit anti-TdT (10 μ l, Lot A02; 30 μ g ml⁻¹ antibody, 100 μ g ml⁻¹ BSA) was applied and incubated for 30 min at room temperature in a humidified chamber. Slides were washed three times in excess PBS and incubated a second time with 10 μ l of FITC-F(ab')₂ anti-rabbit IgG (10 μ g ml⁻¹ in PBS with 100 μ g ml⁻¹ of BSA). Slides were finally washed in excess PBS (0.05% NaN₃) for 6–10 h, mounted in Tris-buffered glycerol, pH 9.5, and counted on a Leitz Ortholux epifluorescent microscope. Over 1,000 consecutive cells were counted from each of five individual mice; values are mean $\pm$ s.d.

It is intriguing to speculate that a failure in production of TdT⁺ lymphoid precursor cells may be an important defect in autoimmune disease, but other animal models of immune dysfunction need to be examined for changes in TdT-bearing cell populations during the progression of disease. Immunodeficient nude mice have near normal bone marrow TdT⁺ populations that develop early in neonatal life¹⁷ in the virtual absence of TdT⁺ thymocytes. The *nu/nu* mouse defect is primarily related to T cells whereas the *me/me* mutation affects both T and B functions.

This work was supported by USPHS grants CA-20823, CA-23262 and NIH grant DE-04887.

Received 9 December 1980; accepted 19 January 1981.

- Green, M. C. & Shultz, L. D. *J. Hered.* **66**, 250 (1975).
- Kincade, P. W. in *Immune Defects of Laboratory Animals* (eds Gershwin, M. E. & Merchant, B.) (Plenum, New York, in the press).
- Shultz, L. D. & Green, M. C. *J. Immun.* **116**, 936 (1976).
- Shultz, L. D. & Zurier, R. B. in *Genetic Control of Autoimmune Disease* (eds Rose, N. R., Bigazzi, P. E. & Warner, N. L.) 229 (Elsevier, New York, 1978).
- Davidson, W. R., Morse, H. C., Sharrow, S. O. & Chused, T. M. *J. Immun.* **122**, 884 (1979).
- Bollum, F. J. *Blood* **54**, 1203 (1979).
- Baltimore, D. *Nature* **248**, 409 (1974).
- Silverstone, A., Cantor, H., Goldstein, G. & Baltimore, D. *J. exp. Med.* **144**, 543 (1976).
- Janossy, G. *et al. J. Immun.* **123**, 1525 (1979).
- Landreth, K. S., Rosse, C. & Clagett, J. *Fedn Proc.* **39**, 806 (1980).
- Goldschneider, I., Metcalf, D., Mandel, T. & Bollum, F. J. *J. exp. Med.* **152**, 438 (1980).
- Goldschneider, I., Gregoire, K. E., Barton, R. W. & Bollum, F. J. *Proc. natn. Acad. Sci. U.S.A.* **74**, 734 (1977).
- Shultz, L. D., Sidman, C. L. & Unanue, E. R. in *Animal Models of Comparative and Developmental Aspects of Immunity and Disease* (eds Gershwin, M. E. & Cooper, E. L.) 260 (Pergamon, New York, 1979).
- Landreth, K. S., Rosse, C. & Clagett, J. (in preparation).
- Hutton, J. J. & Bollum, F. J. *Nucleic Acids Res.* **4**, 457 (1977).
- Sidman, C. L., Shultz, L. D. & Unanue, E. R. *J. Immun.* **121**, 2392, 2399 (1978).
- Sasaki, R., Bollum, F. J. & Goldschneider, I. *J. Immun.* (in the press).

Similarity of teleocidin B and phorbol ester tumour promoters in effects on membrane receptors

Kazuo Umezawa*, I. Bernard Weinstein†, Ann Horowitz‡, Hirota Fujiki‡, Taijiro Matsushima* & Takashi Sugimura‡

* Institute of Medical Sciences, University of Tokyo, Tokyo-108, Japan

† Cancer Center/Institute of Cancer Research, Columbia University College of Physicians and Surgeons, New York 10032, USA

‡ National Cancer Center Research Institute, Tokyo 104, Japan

The tumour promoter 12-*O*-tetradecanoyl-phorbol-13-acetate (TPA) and structurally related phorbol esters effect changes in avian and mammalian cell cultures that mimic transformation by oncogenic viruses or chemical carcinogens and the inhibition or induction of various types of differentiation (for review see refs 1-3). Unlike initiating carcinogens, which seem to act by binding covalently to cellular DNA, the primary site of action of the phorbol ester tumour promoters seems to be the cell membrane^{2,3}; indeed, specific high-affinity saturable receptors for phorbol esters have been identified in cell membranes^{4,5}. The recently discovered class of tumour promoters, the teleocidins⁶, are as potent as TPA in the induction of ornithine decarboxylase in mouse skin, the inhibition of differentiation of Friend erythroleukaemia cells, the induction of HL-60 cell adhesion⁶ and the promotion of tumours on mouse skin⁷. As the teleocidins are structurally unrelated to the phorbol esters, we set out to determine their effects on cell membranes and receptors. We found that in rodent cell cultures, teleocidin B and dihydroteleocidin B have effects similar to those of TPA and that, at nanomolar concentrations, teleocidin inhibits the binding of phorbol esters to cell-surface receptors, which suggests that the action of both classes of compounds may be mediated by the same or a similar receptor system.

TPA enhances the transformation of the mouse embryo cell line C3H 10T1/2 cells on first exposure to radiation or chemical carcinogens⁸⁻¹⁰. It was of interest, therefore, to see if TPA and the teleocidins exert similar biochemical effects in these cells.

An early membrane response to TPA in mouse skin and certain cell cultures is induction of arachidonic acid release from membrane phospholipids, associated with the synthesis of the arachidonic acid products prostaglandin E₁ and F_{2α} (refs 11-13). In Fig. 1, C3H 10T1/2 cells were prelabelled with ³H-arachidonic acid in conditions in which >95% of the acid is incorporated in phospholipids¹². The subsequent addition of TPA, teleocidin B or dihydroteleocidin B (100 ng ml⁻¹) to these cultures led to the rapid release of ³H-arachidonic acid into the media (Fig. 1a), reflecting increased phospholipid degradation. As with TPA, the teleocidin-induced release of arachidonic acid was also associated with increased synthesis of prostaglandins E₁ and F_{2α} (data not shown here). Concentration curve studies indicated that the optimum concentrations of teleocidin B and dihydroteleocidin B were ~10 ng ml⁻¹ and that these agents were actually more potent than TPA in inducing arachidonic acid release (Fig. 1b). The exposure of avian and mammalian cells to TPA causes a rapid enhancement in the uptake of 2-deoxyglucose^{14,15} and certain other nutrients^{2,3}; we found that teleocidin B and dihydroteleocidin B stimulated 2-deoxyglucose uptake by C3H 10T1/2 cells to levels comparable with those obtained for TPA when all three agents were tested at 100 ng ml⁻¹ concentrations (data not shown).

An additional early and specific response of cells to TPA is inhibition of binding of epidermal growth factor (EGF) to its cell-surface receptors¹⁶⁻¹⁹. Although this does not seem to be due to direct binding of TPA to the receptor¹⁷⁻¹⁹, it does provide a highly sensitive marker of membrane changes induced by the

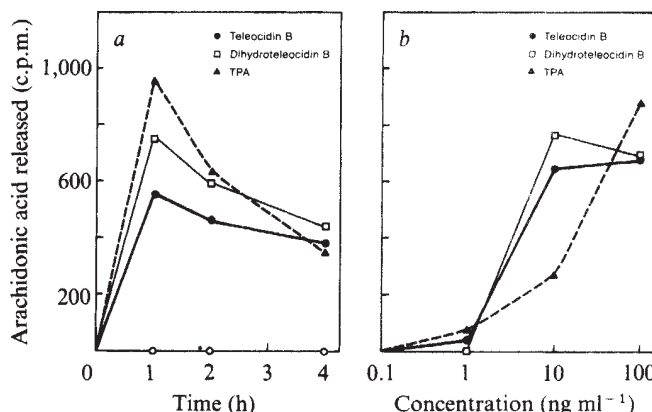


Fig. 1 Arachidonic acid and prostaglandin release were assayed as described elsewhere. Phospholipids of C3H 10T1/2 cells were prelabelled by incubation of 4×10^5 cells per plate with 1.25 μ Ci per ml of ³H-arachidonic acid (62.2 Ci mmol⁻¹, NEN) for 24 h. The prelabelled cells were then washed and incubated in 2 ml of serum-free Dulbecco's minimal essential medium (DMEM) containing either solvent or the indicated chemicals at 37 °C. At the indicated times the media were removed, acidified and extracted with ether. After evaporation of the ether, the sample was solubilized in 0.2 ml of chloroform. A portion (0.04 ml) was applied to silica gel TLC plates with nonradioactive markers, developed in 2,6-dimethylheptanone/glacial acetic acid/0.9% NaCl (80:40:6), and the plate scraped and assayed for radioactivity. Values are means of duplicate samples which agreed within 10%. Two additional experiments gave similar results. *a*, Samples were incubated with 30 ng per ml of TPA (▲), teleocidin B (●), dihydroteleocidin B (□), or the dimethyl sulphoxide (DMSO) solvent (○) for the times indicated. *b*, Samples were incubated with the indicated concentrations of TPA (▲), teleocidin B (●), dihydroteleocidin B (□), or DMSO (○) for 1 h at 37 °C. With the DMSO solvent control there was negligible release of radioactivity, presumably because spontaneous phospholipid turnover is reduced in the serum-free incubation system (ref. 12 and unpublished results).

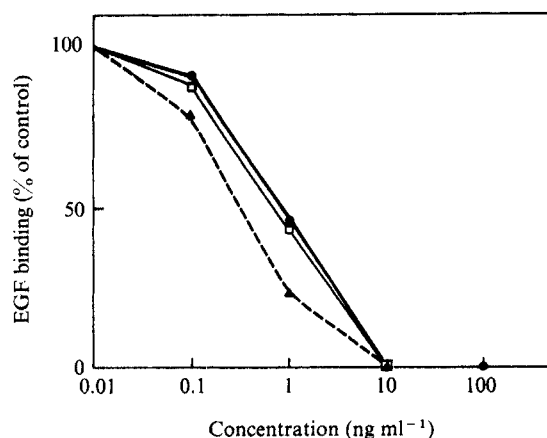


Fig. 2 Inhibition of EGF binding: 2×10^6 C3H 10T1/2 cells were incubated with $0.05 \mu\text{Ci}$ (0.2 ng) of ^{125}I -EGF (120 mCi per mg , Collaborative Research) in 2 ml of serum-free DMEM with the indicated concentration of TPA ($\blacktriangle$), teleocidin B ($\bullet$), or dihydroteleocidin B ($\square$) for 2 h , washed and assayed for cell-bound radioactivity as described elsewhere¹⁷. The control value was $940 \text{ c.p.m. per dish}$ and is expressed as 100% . Values are means of duplicate samples which agreed within 10% . Two additional experiments gave similar results.

phorbol ester tumour promoters. Figure 2 indicates that the addition of 0.1 – 10 ng ml^{-1} of TPA, teleocidin B or dihydroteleocidin B simultaneously with the addition of ^{125}I -EGF led to a marked inhibition of ^{125}I -EGF binding to C3H 10T1/2 cells. All three compounds also inhibited ^{125}I -EGF binding in a line of rat embryo cells (FRE-8D) and in HeLa cells (data not shown). In C3H 10T1/2 and FRE-8D cells, TPA was rather more potent than the teleocidins in inhibiting ^{125}I -EGF binding whereas the three compounds had similar potencies for HeLa cultures.

Several aspects of the action of the phorbol esters have led to the suggestion that they act by usurping cell membrane receptors normally occupied by an endogenous growth factor^{1,16}. Using ^3H -phorbol-12,13-dibutyrate (PDBu) to avoid non-specific binding, Blumberg⁴ has in fact recently demonstrated that membrane fractions from avian and mouse cells contain saturable high-affinity receptors for this compound. A modification of this approach has also demonstrated receptors for ^3H -PDBu on intact monolayers of rat embryo fibroblasts and several other cell types⁵. Figure 3 shows that at 37°C , concentrations of teleocidin B in the range 0.5 – 10 ng ml^{-1} were able to produce marked inhibition of the specific binding of ^3H -PDBu to rat embryo fibroblast monolayer cultures; 50% inhibition was achieved with $\sim 2.5 \text{ ng ml}^{-1}$, which is equivalent to about 10^{-9} M . Teleocidin was more potent than TPA in inhibiting ^3H -PDBu binding, because $\sim 3.8 \text{ ng ml}^{-1}$ of TPA were required to produce 50% inhibition (Fig. 3). The ID_{50} value for PDBu was 23 ng ml^{-1} (data not shown). No inhibition was observed with as much as $1,000 \text{ ng ml}^{-1}$ of phorbol, which lacks tumour promoting activity, nor was inhibition obtained with EGF, presumably because it occupies separate receptors^{4,5,17}. In additional studies (not reported here) of ^3H -PDBu binding, carried out at 4°C rather than 37°C , teleocidin B was also a more potent inhibitor than TPA. Thus, in contrast to the TPA inhibition of EGF binding^{17–19}, teleocidin B inhibition is not highly temperature dependent. These results provide further evidence that teleocidin B competes directly with ^3H -PDBu for receptor binding.

There is evidence that the TPA induction of arachidonic acid release, TPA enhancement of 2-deoxyglucose uptake and TPA inhibition of EGF-receptor binding are direct membrane effects that do not require *de novo* RNA synthesis^{2,11,12,17}. Because of their similar time courses we would assume that this is also true

for the effects of teleocidin on these membrane-related properties. These results, together with our evidence that teleocidin B is a potent and apparently direct inhibitor of the binding of PDBu to its cell-surface receptors, suggest that the teleocidins might act as tumour promoters and induce changes in cell culture by binding to the same membrane-associated receptors as those used by TPA.

We emphasize that although teleocidin B was about 10 times more potent than TPA in inducing arachidonic acid release (Fig. 1) and slightly more potent in inhibiting ^3H -PDBu binding (Fig. 3), it was slightly less potent than TPA in inhibiting EGF-receptor binding (Fig. 2). Because these parameters are quite different and were not all assayed with the same cell cultures, the significance of these quantitative differences is unknown. It is possible that they reflect the existence of subtypes of ^3H -PDBu receptors⁵ that control different cellular effects, but this requires further study.

Teleocidin has an entirely different chemical structure from that of TPA or any of the TPA-related plant diterpenes, but both compounds are somewhat amphipathic. Teleocidin has a peptide-like moiety joined to a lipid portion containing two isoprenes⁶, while TPA contains a somewhat hydrophilic-substituted ring system and a long-chain lipophilic fatty acid residue³. These aspects may provide clues to the molecular details by which the two different classes of compound apparently bind to the same or a similar set of membrane receptors. In addition, the fact that two structurally diverse compounds with potent tumour promoting activity seem to bind to the same cell-surface receptors suggests that these receptors play a part in mediating the biological effects of the compounds.

We thank Drs R. A. Mufson and P. B. Fisher for their advice and Dr M. Takashima for providing dihydroteleocidin B. This research was supported by NCI grant CA-26056, DHEW, an

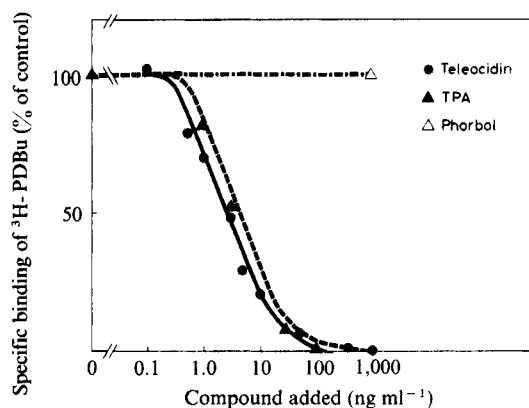


Fig. 3 Inhibition of specific binding of ^3H -PDBu. Rat embryo fibroblast cells (FRE-8D) were grown in 6-cm dishes containing DMEM plus 10% calf serum. The assay was performed on 2×10^6 cells per dish. The cell monolayer was washed once with phosphate-buffered saline (PBS) and then 2 ml of assay buffer (2 volumes DMEM/ 1 volume PBS, plus 1 mg ml^{-1} bovine serum albumin) were added. The indicated concentrations of teleocidin B ($\bullet$), TPA ($\blacktriangle$) or phorbol ($\triangle$) were added, followed by ^3H -PDBu (6.4 Ci mmol^{-1} , Lifesystems) to a final concentration of 3 nM . The cells were incubated for 30 min at 37°C to achieve plateau binding, washed 3 times with assay buffer at 0°C , solubilized in 1 ml of 0.8% Triton X-100, 0.02% EDTA and 0.25% trypsin in PBS, and assayed for radioactivity. Nonspecific ^3H -PDBu binding was assumed to be that obtained in the presence of $50 \mu\text{M}$ unlabelled PDBu. This value was $\sim 20\%$ of the total binding obtained in the absence of excess PDBu and was subtracted from all experimental data to obtain the values for specific saturable binding. Specific binding of ^3H -PDBu in the control assays was $\sim 750 \text{ c.p.m. per } 10^6$ cells (equivalent to $\sim 1.1 \times 10^5$ receptors per cell), with a s.e.m. of $<10\%$. This control value is expressed as 100% . Similar results were obtained in two additional experiments. Further details of this assay will be published elsewhere⁵.

award from the DuPont Company and in part by grants-in-aid for cancer research from the Ministry of Education, Science and Culture and the Ministry of Health and Welfare of Japan.

Received 23 July 1980; accepted 26 January 1981.

- Weinstein, I. B., Wigler, M. & Pietropaolo, C. in *Origins of Human Cancer* (eds Hiatt, H. H., Watson, J. D. & Winston, J. A.) 751–772 (Cold Spring Harbor Laboratory, New York, 1977).
- Weinstein, I. B., Lee, L. S., Fisher, P. B., Mufson, R. A. & Yamasaki, H. *J. supramolec. Struct.* **12**, 195–208, 1979.
- Slaga, T. J., Sivak, A. & Boutwell, R. K. (eds) *Carcinogenesis* Vol. 2 (Raven, New York, 1978).
- Delclos, K. B., Nagle, D. S. & Blumberg, P. D. *Cell* **19**, 1025–1032 (1980).
- Horowitz, A., Greenebaum, E. & Weinstein, I. B. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Fujiki, H., Mori, M., Nakayasu, M., Terada, M. & Sugimura, T. *Biochem. biophys. Res. Commun.* **90**, 976–983 (1979).
- Fujiki, H. *et al.* (in preparation).
- Mondal, S., Brankow, D. W. & Heidelberger, C. *Cancer Res.* **36**, 2254–2260 (1976).
- Mondal, S. & Heidelberger, C. *Nature* **260**, 710–711 (1976).
- Kennedy, A., Mondal, S., Heidelberger, C. & Little, J. B. *Cancer Res.* **38**, 439–443 (1978).
- Levine, L. & Hassid, A. *Biochem. biophys. Res. Commun.* **79**, 477–483 (1977).
- Mufson, R. A., DeFeo, D. & Weinstein, I. B. *Molec. Pharmac.* **16**, 569–578 (1979).
- Yamasaki, H., Mufson, R. A. & Weinstein, I. B. *Biochem. biophys. Res. Commun.* **89**, 1018–1025 (1979).
- Driedger, P. E. & Blumberg, P. M. *Cancer Res.* **37**, 3257–3265 (1977).
- Lee, L. S. & Weinstein, I. B. *J. cell. Physiol.* **99**, 451–460 (1979).
- Lee, L. S. & Weinstein, I. B. *Science* **202**, 313–315 (1978).
- Lee, L. S. & Weinstein, I. B. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5168–5172 (1979).
- Shoyab, M., DeLarco, J. E. & Todaro, G. J. *Nature* **279**, 387–391 (1979).
- Brown, K. D., Dicker, P. & Rozengurt, E. *Biochem. biophys. Res. Commun.* **86**, 1037–1043 (1979).

Synaptic facilitation requires paired activation of convergent pathways in the neocortex

A. Baranyi & O. Fehér

Department of Comparative Physiology, Attila József University of Sciences, Középfasor 52, Szeged 6726, Hungary

In associative learning, the activated neurones undergo a variety of concomitant functional alterations—increases or decreases of firing activity^{1–4} and modifications of membrane potential or resistance^{5–8} and of synaptic responsiveness^{9–11}. Synaptic transmission which can be strengthened only when there is paired activity in two pathways is of particular interest in relation to mechanisms for associative learning^{12,13}. For the neocortex, there are few observations of the plastic changes, induced by conditioning procedures, in the effectiveness of individual synapses^{14–17}. We now report that various regimes with joint stimulations of convergent excitatory pathways on to intracellularly recorded neurones in the motor cortex of the cat result in synaptic facilitation lasting for up to 30 min.

The experiments were performed on the exposed motor cortex of cats anaesthetized with 35 mg per kg Nembutal using conventional intracellular recording techniques. The surgical procedure, methods of stimulation and recording are reported elsewhere in detail^{15,16}. Synaptic responses were elicited by stimulation of the pyramidal tract, the ventrolateral thalamic nucleus (VL), the homotopic cortical surface (callosal stimuli) and the contralateral radial nerve (somatosensory activation). The membrane input resistance, R_m , was measured by the bridge-balancing method¹⁸. At least two excitatory synaptic inputs on to the intracellularly recorded cell were stimulated: one (test pathway) elicited only mono- or oligosynaptic excitatory postsynaptic potentials (test e.p.s.ps), while the other caused the cell to produce spike potentials (unconditioned stimulus, US) (Fig. 1a). Based on the cellular analogues of classical conditioning^{9,19} the applied conditioning paradigm consisted of (1) habituation, (2) pseudoconditioning, (3) conditioning, and (4) extinction procedures (Fig. 1b). A neurone was regarded as conditionable if the test e.p.s.p. was enhanced in

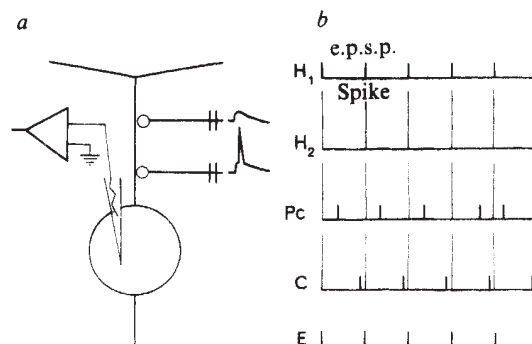


Fig. 1 a, Schematic drawing of the experimental arrangement used for cellular conditioning of neurones in the motor cortex. The stimulation of the test pathway produced an e.p.s.p. while stronger stimuli via another pathway elicited spike potentials in the recorded neurone. b, Illustration of the experimental paradigm. During the habituation procedure 150 test e.p.s.ps (H_1) and then 150 US spikes (H_2) were evoked from the respective inputs at a stimulus frequency of 0.2–1 Hz. PC, pseudoconditioning procedure when test e.p.s.ps and US spikes were randomized with 0.2–1 Hz frequency within a block of 150 stimuli. C, conditioning procedure with e.p.s.p.–spike stimulus pairs at constant inter-stimulus interval ranging from 0 to 100 ms. E, extinction period when only test stimuli were given at 0.2–1 Hz until the augmented neuronal responses returned to the preconditioning level.

amplitude or in efficacy in evoking spike potentials by the conditioning procedure and if these changes were abolished by the extinction procedure.

In 38 cells (26% of the neurones tested) the efficacy of the test e.p.s.ps was enhanced after 60–90 stimulus pairings and the facilitated synaptic response persisted for periods from 3 to 41 min (mean $\pm$ s.d. 14.5 ± 9.5 min, $n = 38$). Figure 2 shows an example of synaptic facilitation by VL e.p.s.p.–callosal spike repetitive stimulus pairs. In other neurones e.p.s.p.–spike and spike–e.p.s.p. sequences were equally efficient, but only in stimulus pairs with inter-stimulus intervals of <40–60 ms in the former and ~100 ms in the latter. Inter-stimulus intervals above 100 ms (periods ranging from 100 to 500 ms were tested) proved ineffective in causing synaptic facilitation. This may be explained by supposing that both US spikes and test e.p.s.ps have an after-effect of about 100 ms in duration and their repetitive overlap is necessary for the facilitatory interaction between them. On the other hand, it could be demonstrated that the observed synaptic facilitation was specific to the pairing procedure, as repeated but unpaired presentation of test e.p.s.ps and US spikes during pseudoconditioning procedures never led to plastic changes (tested in 17 cases, Fig. 2b), and when a previously efficient e.p.s.p.–spike stimulus pair was applied together with a hyperpolarizing current pulse through the recording electrode (when spike generation by US stimuli was prevented), the test e.p.s.p. was not augmented (tested in 13 cases). The amount of hyperpolarization required to prevent action potentials by US stimuli during conditioning was sufficient to abolish the plastic after-effect. Based on the presented data, one may conclude that contingent pairing of the test e.p.s.ps with spike potentials, that is, temporal contiguity of pre- and postsynaptic excitations, is indispensable for the synaptic facilitation, as postulated previously^{20,21}.

Transient facilitations of the test e.p.s.ps lasting for 6–40 s were observed in all 17 neurones with which spike trains of 5–50 Hz frequency and 5–15 s duration were applied as US. This after-effect did not require co-active stimulation of the test and US pathways, because it also occurred when US stimuli were applied on their own. These short-lived effects presumably originated from transient modifications of the ionic micro-environment caused by frequent firing activity^{22–24}. In hippocampal neurones long-term facilitation of both e.p.s.ps and inhibitory postsynaptic potentials has been demonstrated as a

result of repetitive activation with similar parameters^{25,26}. In the motor cortex a considerably higher frequency and longer period of repetitive activation is necessary for long-term facilitation²⁷.

Enhancement of the amplitude of paired test e.p.s.ps was the only consequence of paired conditioning procedures in 21 cells, whereas in the other 17 neurones the synaptic facilitation was accompanied by 3–6 mV membrane depolarization and increases of R_m (Fig. 3). The depolarization presumably did not contribute to the augmentation of the test e.p.s.ps, as a test e.p.s.p. induced during a depolarizing current injection (causing 2–6 mV depolarization) was found in 11 tested cells to be decreased in amplitude or unaffected.

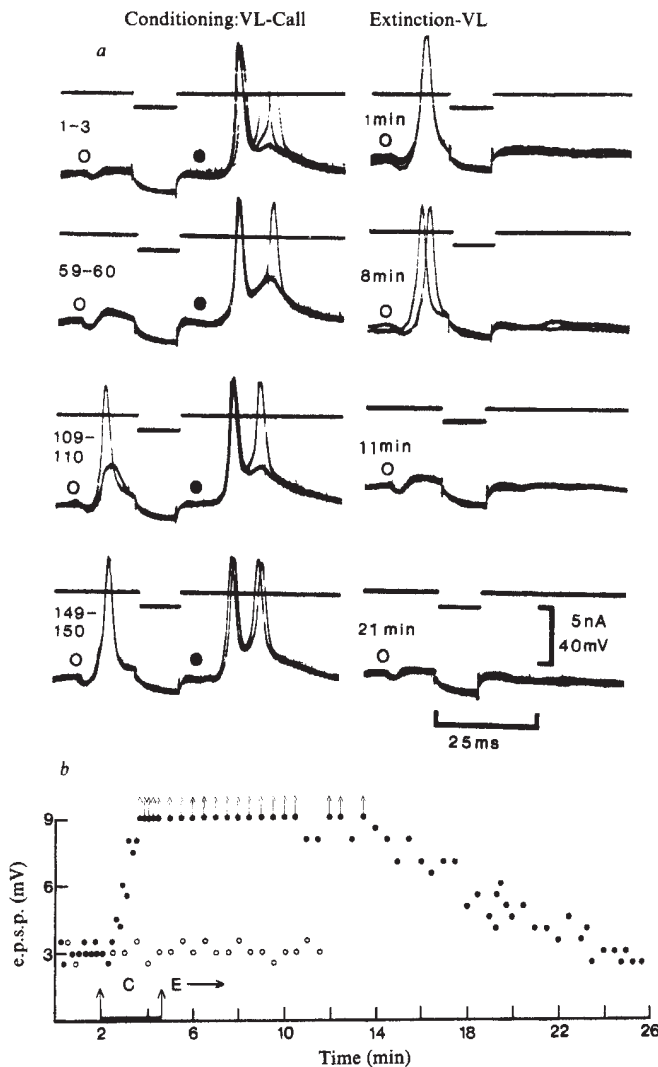


Fig. 2 Facilitation of a test e.p.s.p. by stimulus pairs applied repetitively in a non-pyramidal tract neurone²⁹. Top traces, current passed through the recording microelectrode; bottom traces, intracellular records. The test e.p.s.ps were evoked by stimulation of the VL nucleus (○); US spikes were elicited from the callosal system (Call, ●). The VL e.p.s.p. had a latency of 1.6 ms and was monosynaptic being able to follow 100 Hz frequency of stimulation. *a*, Neuronal responses during the conditioning and extinction procedures repeated once every second with 27 ms inter-stimulus interval. Following the 110th pairing all the test stimuli elicited spike potentials. The enhanced synaptic response returned to the preconditioning level by about the 21st min of the extinction period. R_m was measured every second with 1.2 nA anodal current pulses. No changes in R_m or spontaneous firing were observed during conditioning. *b*, Time course and specificity of facilitation of test e.p.s.p. for pairing procedure. ● (Paired e.p.s.p.) represents an actual data point every 10th second of the conditioning and every half minute of the extinction period. In the unpaired situation (○), only data at each half-minute intervals are presented.

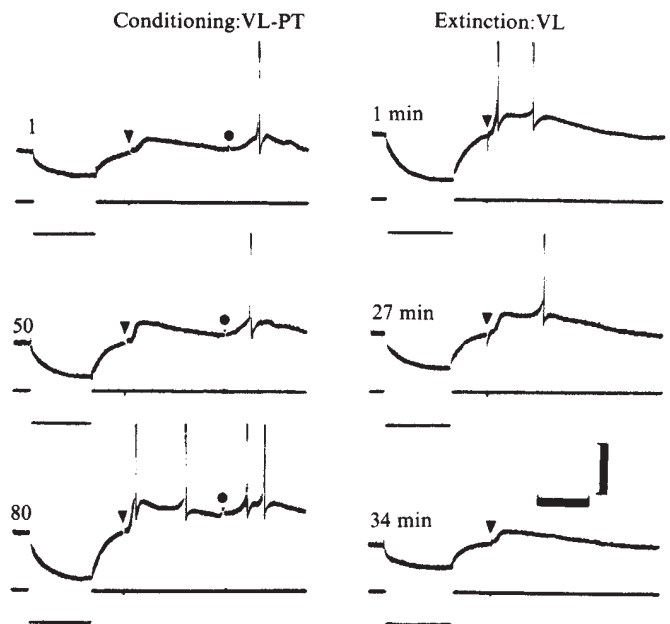


Fig. 3 Conditioned facilitation of a VL test e.p.s.p. (▼) in a non-pyramidal tract neurone with accompanying membrane depolarization and increases in R_m . The US spikes were elicited by pyramidal tract (PT) stimuli (●). Inter-stimulus interval 30 ms, frequency of pairing 0.5 Hz. Top traces, intracellular records; bottom traces, current records. Only the lower part of action potentials are shown because of high amplification. Calibration: 15 mV for intracellular records and 1.66 nA for current records. Time calibration 15 ms. The amplitude of PT spikes was 71 mV. The test e.p.s.ps and PT spikes were monosynaptically evoked (latency, 1.2 and 4.8 ms at 100 Hz stimulus frequency, respectively). During conditioning the preconditioning value of the amplitude of test e.p.s.ps (mean $\pm$ s.d. 4 ± 0.2 mV, $n = 10$) gradually increased and attained maximal facilitation by the 80th pairing, when each VL stimulus evoked two spikes. Note the conditioned decreases of the mean time to peak of VL e.p.s.ps from 2.7 ± 0.1 to 1.7 ± 0.1 ms ($n = 5$). Parallel with the synaptic facilitation, the R_m increased from 5.1 ± 0.3 to 8.8 ± 0.3 M Ω ($n = 10$), while the membrane potential depolarized from 67 ± 2 to 63 ± 2 mV ($n = 10$). All the conditioned changes of test e.p.s.ps and membrane parameters completely disappeared by the 34th min of the extinction period, when only VL stimuli were given alone with 0.5 Hz frequency. None of the above changes occurred during the pseudo-conditioning procedure, when PT and VL stimuli were randomized.

Intracellular observations in the motor cortex of the cat²⁸ and in invertebrate neurones⁵ have suggested that in some conditions an inward calcium current is initiated which leads, via a second messenger system, to membrane depolarization and an elevation in R_m . On the other hand, the R_m increases associated with conditioned increases in firing rate might be due to a decrease in K^+ conductance⁸. Our observations demonstrate that conditioned increases of synaptic efficacy originate from some kind of, as yet unknown, focal postsynaptic processes in the motor cortex without any general modification of the neuronal excitability. Although concomitant alterations of membrane potential or conductance did occur in some of the conditioned neurones, presumably they represent a different type of plastic modification. The mechanism of synaptic facilitation described here requires further examination.

Received 8 April 1980; accepted 28 January 1981.

- Kandel, E. R. & Spencer, W. A. *Physiol Rev.* **48**, 65–134 (1968).
- Berger, T. W. & Thompson, R. F. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1572–1576 (1978).
- Morrel, F. in *The Neurosciences, A Study Program* (eds Quarten, G. C., Melnechuk, T. & Schmitt, F. O.) 452–469 (Rockefeller University Press, New York, 1967).
- Morrel, L. T., Hoepfner, T. J. & Morrel, F. *Science* **204**, 528–530 (1979).
- Alkon, D. L. *Science* **205**, 810–816 (1979).
- Woody, C. D. & Black-Cleworth, P. J. *Neurophysiol.* **36**, 1104–1115 (1973).

7. Kandel, E. R. & Tauc, L. J. *J. Physiol., Lond.* **183**, 287–304 (1966).
8. Hoyle, G. *Neurosci. Res. Prog. Bull.* **17**, No. 4, 577–586 (1979).
9. Kandel, E. R. & Tauc, L. J. *J. Physiol., Lond.* **181**, 1–47 (1965).
10. Kandel, E. R. & Tauc, L. *Nature* **202**, 145–147 (1964).
11. Voronin, L. L. in *Soviet Research Reports Vol. 2* (ed. Woody, C. D.) (Brain Information Service, University of California, Los Angeles, 1976).
12. McNaughton, B. L., Douglas, R. M. & Goddard, G. V. *Brain Res.* **157**, 277–293 (1978).
13. Gardner-Medwin, A. R. *Trans. Biochem. Soc.* **6**, 841–844 (1978).
14. Voronin, L. L. *Acta neurobiol. exp.* **40**, 1B, 335–370 (1980).
15. Baranyi, A. & Fehér, O. *Expl Brain Res.* **33**, 283–298 (1978).
16. Baranyi, A. & Fehér, O. *Expl Brain Res.* **41**, 124–134 (1981).
17. Tzebelikos, E. & Woody, C. D. *Soc. Neurosci. Abstr.* **4**, 81 (1978).
18. Nelson, P. G. & Frank, K. J. *Neurophysiol.* **30**, 1097–1113 (1967).
19. Kandel, E. R. in *The Neurosciences, A Study Program* (eds Quarten, G. C., Melnechuk, T. & Schmitt, F. O.) 666–689 (Rockefeller University Press, New York, 1967).
20. Hebb, D. O. *The Organization of Behaviour* (Wiley, New York, 1949).
21. Konorski, J. *Conditioned Reflex and Neurone Organization* (Cambridge University Press, London, 1948).
22. Heinemann, U. & Lux, H. D. *Brain Res.* **93**, 63–76 (1975).
23. Koike, H., Mano, N., Okada, Y. & Oshima, T. *Expl Brain Res.* **14**, (1972).
24. Moody, W., Futamachi, K. & Prince, D. A. *Expl Neurol.* **42**, 248–263 (1974).
25. Misgeld, U., Sarvey, J. M. & Klee, M. R. *Expl Brain Res.* **37**, 217–229 (1979).
26. Lynch, G., Gall, C. & Dunwiddie, T. *Prog. Brain Res.* **48**, 113–130 (1978).
27. Bindman, L. J., Lippold, O. C. J. & Milne, A. R. *J. Physiol., Lond.* **286**, 457–477 (1979).
28. Woody, C. D., Swartz, B. E. & Gruen, E. *Brain Res.* **158**, 373–375 (1978).
29. Takahashi, E. *J. Neurophysiol.* **27**, 908–924 (1965).

Response of an insect photoreceptor: a simple log-normal model

Richard Payne & Jonathon Howard

Department of Neurobiology, Research School of Biological Sciences, Australian National University, PO Box 475, Canberra City, ACT 2601, Australia

Photoreceptors respond to light with a graded depolarization or hyperpolarization of the cell membrane. The transduction of light energy to an electrical event, though much studied, remains an outstanding problem in visual research. Previous transduction models¹, involving cascades of chemical reactions, do not fit the responses of insect photoreceptors². It has been proposed that the absorption of light causes the release of an internal transmitter³. We have developed here a simple two-parameter model of transduction based on the assumption that membrane channels are activated when the concentration of internal transmitter reaches a threshold level⁴. The theory predicts that the impulse-response function follows the log-normal curve—this model fits the response of the locust compound eye photoreceptor over the 10-fold range of time scale produced by altering the temperature or the state of light adaptation of the cell.

Photoreceptors in the compound eye of *Locusta migratoria* generate a variable number of superimposed, depolarizing voltage events when stimulated with a brief, weak flash of light⁵. Each event or 'bump' represents an effective quantal absorption of light⁶. The average of a large number of such responses accurately defines the impulse-response function of the receptor. The amplitude of this response increases linearly with increasing flash intensity when the responses have an amplitude of <5 mV and in this linear range, the Fourier transformation of the impulse-response function corresponds to the temporal frequency response of the receptor⁷.

Using glass micropipettes, recordings of electrical potential were made from photoreceptors in an intact eye⁸ and a superfused retina⁹. For the superfused retina we averaged responses from dark-adapted photoreceptors to 1-ms flashes from a green light-emitting diode (Siemens LD57C) at temperatures of the surrounding perfusion fluid between 33 and 17 °C. When recording from photoreceptors in an intact eye, a 1-ms test flash from a point light source (also a LED) could be superimposed on steady, adapting illumination from the same source which was

placed on the optical axis of the ommatidium containing the imaged cell.

We attempted to fit our impulse-response functions to the models of Fuortes and Hodgkin¹ and Baylor *et al.*¹⁰. These models are based on the time course of the concentration of the penultimate product of a series of n reactions which proceed sequentially¹ or independently¹⁰. Like Borsellino *et al.*¹¹, we found that the Fuortes and Hodgkin model predicted a response that was too symmetrical about the peak and it also required a range of values of n between 13 and 20 to fit the data from all our cells. The model of Baylor *et al.* fitted the response shape adequately, but required values of n between 20 and 48. The large values for the number of steps required in the reaction chains of these models and the variability from cell to cell made

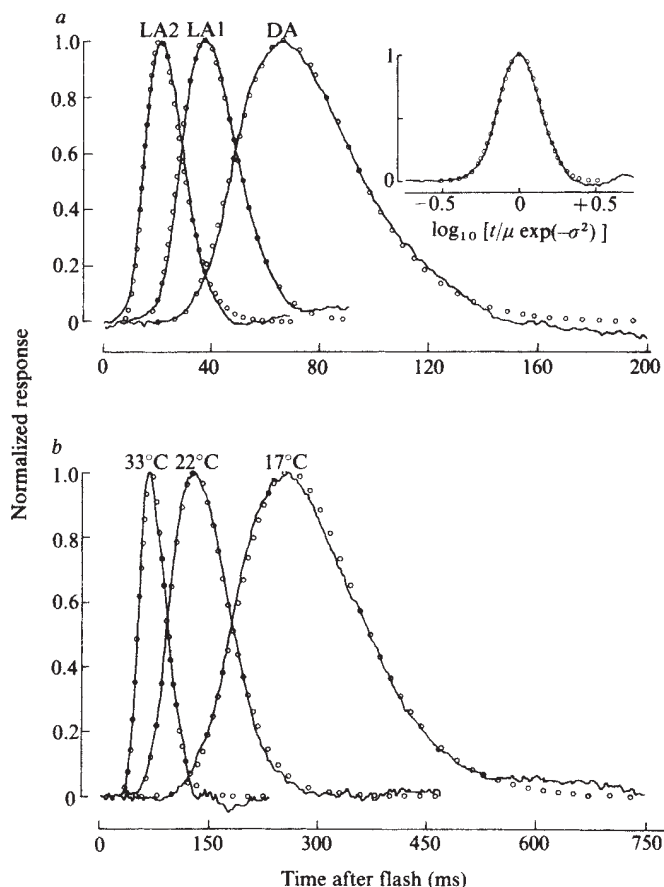


Fig. 1 *a*, Impulse-response functions recorded from a photoreceptor in an intact eye when dark adapted and at two light-adapted states. One hundred presentations of a 1-ms light flash were averaged in each case. The intensities of the flashes and superimposed adapting light were: 14 photons per facet per flash (dark adapted), 42 photons per facet per flash superimposed on an adapting light of 1.3×10^4 photons per facet per s and 4.2×10^2 photons per facet per flash superimposed on an adapting light of 1.3×10^5 photons per facet per s, wavelength 560 nm. Circles show the fits to the data using equation (1) and the following parameters: DA: $\sigma = 0.31$, $\mu = 74$ ms; LA1: $\sigma = 0.27$, $\mu = 41$ ms; LA2: $\sigma = 0.31$, $\mu = 23$ ms. The amplitudes of the responses were 1.5 (DA), 1.2 (LA1) and 2.3 (LA2) mV. Inset, the impulse-response function in the dark-adapted state is replotted on a logarithmic time base. Equation (1) (circles) predicts that a gaussian curve will result, having mean $\mu \exp(-\sigma^2)$ and standard deviation σ . *b*, Impulse-response functions of a photoreceptor in a perfused eye at three temperatures. Circles show the fit to the data of equation (1) using the following parameter values: 17 °C: $\sigma = 0.305$, $\mu = 394$ ms; 22 °C: $\sigma = 0.294$, $\mu = 142$ ms; 33 °C: $\sigma = 0.247$, $\mu = 75$ ms. The responses shown are averages of 100 presentations of a 1-ms flash of intensity 1×10^{12} photons $\text{cm}^{-2} \text{s}^{-1}$ and wavelength 560 nm. The amplitudes of the responses, which have been scaled to unit height, were 3.1 (17 °C), 3.9 (21 °C) and 3.8 (33 °C) mV.

us doubt their physical applicability. We therefore searched for an alternative model.

We chose to fit our impulse-response functions to the two-parameter log-normal curve (Fig. 1a)

$$V(t) = sI \Delta t \cdot t^{-1} \cdot \exp[-(\log(t/\mu))^2/2\sigma^2] \quad (1)$$

where I is the intensity of the flash of duration Δt delivered at $t = 0$ and s is the sensitivity of the cell. σ is the standard deviation and μ the mean of the transformed gaussian curve.

An adapting background of intensity 3×10^4 photons per facet per s reduced the value of μ for responses from cells in the intact eye from 60 ± 18 ms ($n = 6$) in the dark-adapted state to 24 ± 2 ms when light adapted. The corresponding values of σ , 0.30 ± 0.01 and 0.32 ± 0.02 , showed no significant change (for example, see Fig. 1a). A large decrease in the sensitivity of the response also occurred during light adaptation as noted by others^{12,13}. Lowering the temperature of the superfused preparation from 33 to 21 °C resulted in an increase in μ from 70 ± 22 ms ($n = 7$) to 160 ± 14 ms. A marginal increase in σ from 0.28 ± 0.02 to 0.32 ± 0.02 also occurred (for example, Fig. 1b). Temperature and light adaptation therefore significantly affect only one temporal parameter in our model (μ), corresponding to a change in the time scale of the response. We propose that this parameter represents the rate constant of a single transduction process, being sensitive to both light adaptation and temperature.

The log-normal distribution has a general application to the analysis of biological data with a large variance¹⁴. In particular, Gaddum¹⁵ and Bliss¹⁶ have used the distribution to describe the variability, from subject to subject, in the critical dose of a drug that will just cause a reaction. We propose an analogous physical model of phototransduction. The model assumes that, following the absorption of light, the strength of a chemical signal increases. The broad impulse-response arises from a scatter in the detection of threshold levels of this signal that trigger the activation of depolarizing ionic channels in the photoreceptive membrane.

Specifically, we assume that: (1) a photoisomerization opens several ionic channels in the vicinity of the activated rhodopsin molecule through the action of an internal transmitter. A given channel opens for a brief time when the concentration of transmitter exceeds a threshold value, α . (2) The threshold

value, α , is normally distributed from channel to channel. (3) The time course of local transmitter concentration following a photoisomerization is a smooth, monotonic function rising from zero at the instant of the stimulus to a positive value, before declining. Its time course is governed by a single time constant, k , so that the transmitter concentration is $y = f(kt)$, where t is the time elapsed after the light flash. (4) The time constant, k , is chosen so as to minimize the temporal width of the voltage generated for a given time to peak of the response t_p .

A direct consequence of assumptions (3) and (4) is that the transmitter concentration, $f(kt)$, increases almost logarithmically with respect to time in the neighbourhood of t_p . This is because, by assumption (4), k is chosen so as to maximize the rate of increase in transmitter concentration, $\partial f(kt)/\partial t$, at $t = t_p$; so that

$$\frac{\partial^2 f(kt)}{\partial k \partial t} = 0 \quad \text{at } t = t_p \quad (2)$$

However, from the identity

$$\frac{\partial f(kt)}{\partial \log t} = t \frac{\partial f(kt)}{\partial t} = k \frac{\partial f(kt)}{\partial k} \quad (3)$$

it follows that for our optimized system

$$\frac{\partial^2 f(kt)}{(\partial \log t)^2} = kt \frac{\partial^2 f(kt)}{\partial k \partial t} = 0 \quad \text{at } t = t_p \quad (4)$$

Integrating equation (4) twice with respect to $\log(t)$ yields $f(kt) = a \log(t) + b$ in the neighbourhood of t_p (a, b constant). The impulse-response function represents the sum of the voltages generated by all the channels that open following a light flash. This population of channels has, by assumption (2), thresholds to transmitter concentration that are normally distributed. If the concentration of transmitter rises logarithmically with time it follows that the temporal frequency of opening of channels will be log-normal so that, if channel opening times are brief, the voltage generated will follow equation (1). The model is illustrated in Fig. 2.

Our model describes the photoreceptor response as the result of a variable threshold associated with a process that is logarithmically related to time. Of the two temporal parameters only μ , which describes the rate of release of the internal transmitter, is significantly affected by temperature or light adaptation. We intend to continue our investigations by examining the response of 'faster' insect eyes, such as those of the fly and dragonfly. If equation (1) can be applied to other photoreceptors, then the two parameters, μ and σ , will provide a simple method of comparing their temporal frequency response, similar to the parameters describing the approximately gaussian angular sensitivity curves of photoreceptors¹⁷.

Received 2 October 1980; accepted 15 January 1981.

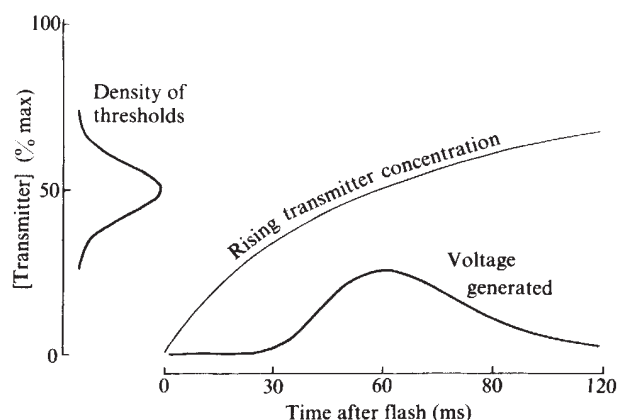


Fig. 2 Illustration of the physical model proposed in the text. The gaussian curve describing the thresholds for the opening of channels is transformed into the asymmetric impulse-response through the non-linearity of the rising 'transmitter' concentration. The Langmuir isotherm, $y = kt/(1 + kt)$ is chosen to describe the transmitter concentration, y . In the region of the activation of channels, between $y = 0.3$ and 0.7 , this equation is approximated by $y = 0.25 \log(t) + c$. Choosing to activate channels in this logarithmic region provides for the least scatter in channel opening times.

1. Fuortes, M. G. F. & Hodgkin, A. L. *J. Physiol., Lond.* **172**, 239–263 (1964).
2. French, A. S. *Nature* **283**, 200–202 (1980).
3. Cone, R. A. in *Biochemistry and Physiology of Visual Pigments* (ed. Langer, H.) 275–282 (Springer, Berlin, 1973).
4. Hamdorf, K. & Kirschfeld, K. *Z. Naturforsch.* **35c**, 173–174 (1980).
5. Scholes, J. *Cold Spring Harb. Symp. quant. Biol.* **30**, 517–527 (1965).
6. Lillywhite, P. G. *J. comp. Physiol.* **122**, 189–200 (1977).
7. Pinter, R. B. *J. comp. Physiol.* **77**, 383–397 (1972).
8. Wilson, M. J. *comp. Physiol.* **97**, 323–328 (1975).
9. Payne, R. *Biophys. struct. Mech.* **6**, 235–251 (1980).
10. Baylor, D. A., Hodgkin, A. L. & Lamb, T. D. *J. Physiol., Lond.* **242**, 685–727 (1974).
11. Borsellino, A., Fuortes, M. G. F. & Smith, T. G. *Cold Spring Harb. Symp. quant. Biol.* **30**, 429–443 (1965).
12. Zettler, F. Z. *vergl. Physiol.* **64**, 432–449 (1969).
13. French, A. S. *Biol. Cybernet.* **32**, 115–123 (1979).
14. Gaddum, J. H. *Nature* **156**, 462–466 (1945).
15. Gaddum, J. H. *MRC spec. Rep.*, 183 (1933).
16. Bliss, C. I. *Ann. appl. Biol.* **24**, 134–166 (1937).
17. Washizu, Y., Burkhardt, D. & Streck, P. Z. *vergl. Physiol.* **48**, 413–428 (1964).

Choline biosynthesis by a preparation enriched in synaptosomes from rat brain

Jan K. Blusztajn & Richard J. Wurtman

Laboratory of Neuroendocrine Regulation, Department of Nutrition and Food Science, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

It is widely held that brain cells are unable to synthesize choline *de novo*, and that the only source of this compound for brain acetylcholine or membrane biosynthesis is the choline or choline-containing phospholipids taken up from the circulation¹. This notion has been difficult to reconcile with observations²⁻⁴ that there is a net efflux of choline from the brain. Recently we⁵ and others^{6,7} have demonstrated that various preparations of mammalian brain contain enzymes, the phosphatidylethanolamine *N*-methyltransferases (PeMT), which catalyse the synthesis of phosphatidylcholine (PC), using *S*-adenosylmethionine (SAM) as a methyl donor for the stepwise methylation of phosphatidylethanolamine (PE). The highest specific activity of PeMT was present in synaptosomal preparation^{5,6}. We now report that rat brain synaptosomal preparations can also metabolize the PC generated by PeMT to liberate free choline.

Male Sprague-Dawley rats (180–220 g, Charles River) were housed under light (Vita-Lite, Duro-Test) for 12 h daily and had free access to water and food (Charles River Rat Chow). Animals were decapitated, their brains quickly dissected and synaptosomal preparations of whole brains (excluding cerebellum) were obtained as described previously⁵; their identity was confirmed by electron microscopy. Synaptosomal preparations were incubated in Krebs-Ringer bicarbonate buffer (pH 7.4) at 37 °C. Heat-treated synaptosomes (100 °C, 20 s) were used to generate blank values. Phosphatidyl-*N,N*-dimethylethanolamine (P2) (Gibco) was used as PeMT substrate, and [*Me*-³H]*S*-adenosylmethionine (10 mCi μ mol⁻¹, NEN) as the methyl donor. Enzymatic reactions were stopped by addition of the appropriate extraction solution. Phospholipids were extracted and purified as described before⁵, and the ³H-methyl present in newly formed ³H-PC was quantified by liquid scintillation spectrophotometry. The choline, including [*Me*-³H]choline, was extracted from the incubation media with 0.5% tetraphenylboron (Aldrich) in heptanone (Aldrich), and back-extracted from the heptanone phase into 0.4 M hydrochloric acid. The aqueous layer was then dried under vacuum, and the dry residue redissolved in ethanol and subjected to

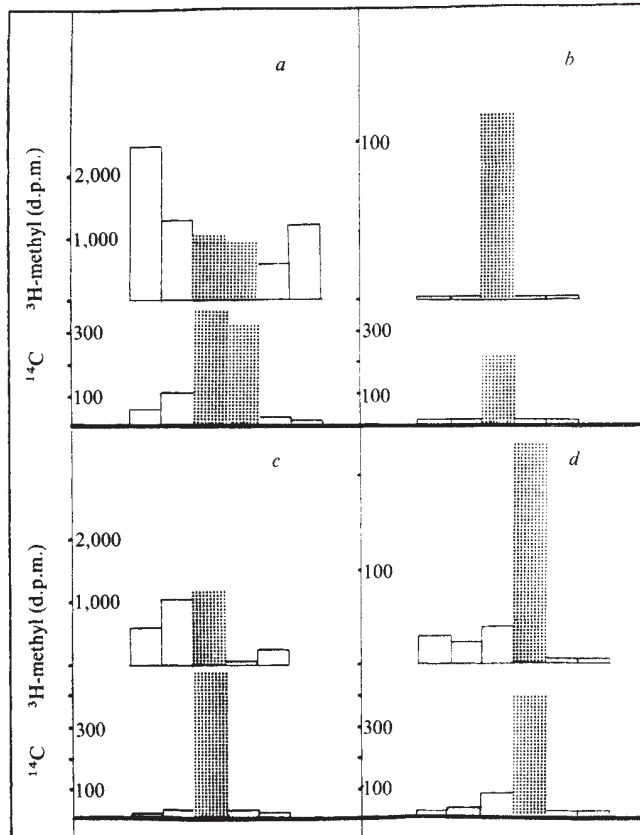


Fig. 1 Purification of newly synthesized choline. Synaptosomal preparations (0.5 mg protein) were incubated at 37 °C for 30 min in 0.1 ml Krebs-Ringer bicarbonate buffer (pH 7.4) containing: 4.5 μ M [*Me*-³H]SAM (10 mCi μ mol⁻¹), 50 μ g of ultrasound-dispersed P2, and 1,700 d.p.m. of [*Me*-¹⁴C]choline (60 mCi mmol⁻¹). The reaction was stopped by addition of 0.3 ml of 0.5% tetraphenylboron in heptanone. The choline in the heptanone was back-extracted to 0.3 ml of 0.4 M HCl, which was then dried under vacuum. The dry residue was then redissolved in 0.05 ml of ethanol and applied to cellulose TLC plates (a) or to filter paper for electrophoresis (c). The TLC solvent system contained: *n*-butanol/ethanol/acetic acid/water 40:10:1:15 (by volume). The electrophoretic conditions were: 0.75 M acetic acid, 0.75 M formic acid pH 2.0, constant voltage of 17 V cm⁻¹ for 40 min. Each sample also contained 5 μ g of choline chloride carrier. The plates and papers were stained for choline¹⁶, and segments staining positively were extracted with 50% ethanol in 0.2 M HCl. The extracts were then concentrated and rechromatographed as described above, except that TLC-derived material (a) was subjected to electrophoresis (b), while electrophoresis-derived material (c) was subjected to TLC (d). a, b, c and d show distributions of radioactivity on chromatograms prepared as described above. Shaded areas indicate segments which stained positively for choline. Lower panels show the distribution of ¹⁴C-choline standards and provide estimates of choline recovery. Incubations using heat-treated synaptosomal preparations yielded flat background radioactivity patterns in b and d.

paper electrophoresis⁸ and TLC. As described below, both procedures were needed to reduce background radioactivity. The radioactivity in segments of the chromatograms was counted by liquid scintillation spectrophotometry. The overall recovery of choline was calculated by adding ¹⁴C-choline as an internal standard to synaptosomal incubations. The identity of the ³H-choline formed enzymatically was established by subjecting the water-soluble products of the methylation reactions, prepurified by TLC, to phosphorylation by choline kinase and purifying the products by paper electrophoresis. Total concentrations of choline and PC in the incubations were assessed by radioenzymatic assay of choline⁹, using choline kinase and [γ -³²P]ATP; the PC was purified by TLC and acid hydrolysed to choline before assay.

Table 1 Phosphorylation of newly synthesized choline by choline kinase

	Choline (d.p.m.)	Phosphorylcholine (d.p.m.)
Control	257 $\pm$ 22	0 (background)
Choline kinase	31 $\pm$ 7	323 $\pm$ 51

Incubations, extractions and TLC prepurification of newly synthesized choline were carried out as in Fig. 1 legend. Choline was extracted from TLC plates with 50% ethanol in 0.2 M HCl and the extracts were then dried under vacuum. The dry residues were resuspended in 0.1 ml of medium containing 12 mM glycylglycine, 8 mM dithiothreitol, 18 mM MgCl₂, 1.2 mM ATP and 0.05 U choline kinase (pH 7.4), and incubated for 15 min at 30 °C. Reaction was stopped with 0.1 ml ethanol and the mixtures concentrated and subjected to paper electrophoresis (as described in Fig. 1 legend). Segments were visualized for choline¹⁶ and phosphorylcholine¹⁷ and then the radioactivity counted. Data are expressed as means $\pm$ s.d.

Identification of newly synthesized choline proved relatively difficult. It was not until we combined TLC and paper electrophoresis that we could see a sharp ^3H -methyl radioactive peak that co-migrated with authentic nonradioactive and ^{14}C -choline (Fig. 1). The heat-treated preparation of synaptosomes produced no radioactive product that would co-migrate with authentic choline in either of these chromatographic systems. After subjecting the ^3H -methyl material that had co-chromatographed with authentic choline to phosphorylation by choline kinase, virtually none of the radioactivity co-chromatographed with choline (Table 1). The radioactivity could, however, be recovered from material co-migrating with phosphorylcholine, as would be expected if the ^3H -methyl material were authentic choline.

The rate of incorporation of ^3H -methyl groups into choline depended on the concentration of PeMT substrate present—addition of increasing amounts of P2 to incubation mixtures (Fig. 2) caused parallel increases in the apparent rates of synthesis of PC and choline until saturation occurred at $\sim 500\ \mu\text{g}$ P2 per ml. This increase in apparent synthesis rate was more pronounced for ^3H -PC, where the ratio of rate of ^3H -methyl incorporation when 1 mg per ml of P2 was present to incorporation without added P2 was 5.7; the same ratio for ^3H -choline was only 1.5. Addition of dimethylethanolamine (30 mM final concentration) did not enhance the synaptosomal formation of choline (data not shown). These results indicate that the ^3H -choline must have been formed by degradation of newly synthesized PC, or, less likely, by action of some as yet undiscovered enzyme on a degradation product of P2.

The SAM concentration at which our assay was carried out ($4.5\ \mu\text{M}$) is below the apparent K_m for PeMT^{5,6}. However, at this SAM concentration the reaction proceeded linearly with time for up to 50 min. The incorporation of ^3H -methyl into ^3H -PC proceeded at the rate of 1.3 pmol per mg protein per 30 min, and into ^3H -choline at 0.4 pmol per mg protein per 30 min (Table 2). The choline newly formed during this interval would represent 88 p.p.m. of the total choline in the synaptosomal preparation, whereas the newly formed PC would represent only 1.9 p.p.m. of total PC (Table 2). Thus the relative enrichment of the choline pool was 47-fold greater than that of PC. These results suggest that there may exist specific physical domains of phospholipids within synaptosomal membranes, such that the PE methylation and PC degradation occur in the same domain. This would explain why up to 23% of the [Me - ^3H]choline present in the incubations could be recovered as free choline molecules and not covalently bound in the form of PC, while free choline constituted $<1\%$ of the total amount of

Table 2 Comparison of synthesis of phosphatidylcholine and choline by rat brain synaptosomal preparations

	Phosphatidylcholine	Choline
^3H -methyl incorporation (pmol per mg protein per 30 min)	1.323 ± 0.219	0.396 ± 0.086
Concentration (nmol per mg protein)	708 ± 29	4.5 ± 0.82
Specific activity (p.p.m.)	1.87 ± 0.37	88 ± 34
I_{Ch}		
I_{PC}	0.30	
SA_{Ch}		
SA_{PC}	47.3	

Incubation, extraction and purification conditions were as described in Fig. 1 and Table 1 legends. Free ^3H -choline and ^3H -choline hydrolysed off ^3H -PC were determined by radioenzymatic assay⁹. Data are expressed as means $\pm$ s.d. ^3H -methyl incorporations are I_{PC} and I_{Ch} , specific activities are SA_{PC} and SA_{Ch} for phosphatidylcholine and choline, respectively.

choline (that is free plus PC-bound) in the synaptosomal preparations (Table 2). Other studies have suggested the existence of two PC pools in synaptosomes, one with a $t_{1/2}$ of 2 days and another with $t_{1/2}$ of 52.5 days (ref. 10). Our results indicate the existence of yet another pool, which turns over much more rapidly, and which contains PC molecules derived from the PeMT pathway.

The brain is reported to contain several enzyme systems able to catalyse the degradation of PC to liberate free choline: phospholipases A_1 and A_2 and lysophospholipase¹¹ generate glycerol-3-phosphorylcholine, which can subsequently be hydrolysed in a one-step¹² or two-step process¹³ to yield free choline; phospholipase D can catalyse PC hydrolysis to liberate free choline directly¹⁴; lysophospholipase D could liberate choline from lysophosphatidylcholine formed by the action of phospholipase A_2 (ref. 15). Our data do not indicate which, if any, of these enzyme systems liberate ^3H -choline from ^3H -PC formed in synaptosomal preparation by the transmethylation pathway—our present techniques are insufficient to detect other [Me - ^3H]choline containing species, which might provide clues.

It seems clear that previously, by measuring only phospholipid products, we⁵ and others^{6,7} have underestimated the contribution of PeMT to brain choline. It is entirely possible that other compounds, such as glycerol-3-phosphorylcholine and phosphorylcholine, also accumulate in our experimental system as intermediates in the genesis of free choline; thus we may still be underestimating the total rate of choline production in the brain. The absolute contribution of *de novo* synthesis to total brain choline and the extent to which this source provides precursor molecules for acetylcholine formation await further study.

These studies were supported in part by a grant from the National Institute of Mental Health. We thank Mr Robert Scheinman for technical assistance.

Received 15 September 1980; accepted 20 January 1981.

1. Ansell, G. B. & Spanner, S. in *Cholinergic Mechanisms and Psychopharmacology* (ed. Jenden, D. J.) (Plenum, New York, 1977).
2. Aquilonius, S. M., Ceder, G., Lying-Tunell, U., Malmund, H. O. & Schubert, J. *Brain Res.* **99**, 422–430 (1975).
3. Choi, R. L., Freeman, J. J. & Jenden, D. J. *J. Neurochem.* **24**, 735–741 (1975).
4. Dross, K. & Kewitz, H. N.-S. *Archs Pharmac.* **275**, 91–106 (1972).
5. Blusztajn, J. K., Zeisel, S. H. & Wurtman, R. J. *Brain Res.* **179**, 319–327 (1979).
6. Crews, F. T., Hirata, F. & Axelrod, J. *J. Neurochem.* **34**, 1491–1498 (1980).
7. Mozzi, R. & Porcellati, G. *FEBS Lett.* **100**, 363–366 (1979).
8. Potter, L. T. & Murphy, W. *Biochem. Pharmac.* **16**, 1386–1388 (1967).
9. Goldberg, A. M. & McCaman, R. E. *J. Neurochem.* **20**, 1–8 (1973).
10. Pasquini, J. M., Krawiec, L. & Soto, E. F. *J. Neurochem.* **21**, 647–653 (1973).
11. Cooper, M. F. & Webster, G. R. *J. Neurochem.* **17**, 1543–1554 (1970).
12. Webster, G. R., Marples, E. A. & Thompson, R. H. S. *Biochem. J.* **65**, 374–377 (1975).
13. Abra, R. M. & Quinn, P. J. *Biochim. biophys. Acta* **380**, 436–441 (1975).
14. Saito, M. & Kanfer, J. *Archs Biochem. Biophys.* **169**, 318–323 (1975).
15. Wykle, R. L. & Schremmer, J. M. *J. biol. Chem.* **249**, 1742–1746 (1974).
16. Beiss, U. J. *Chromatography* **13**, 104–110 (1964).
17. Wade, H. E. & Morgan, D. M. *Nature* **171**, 529–530 (1953).

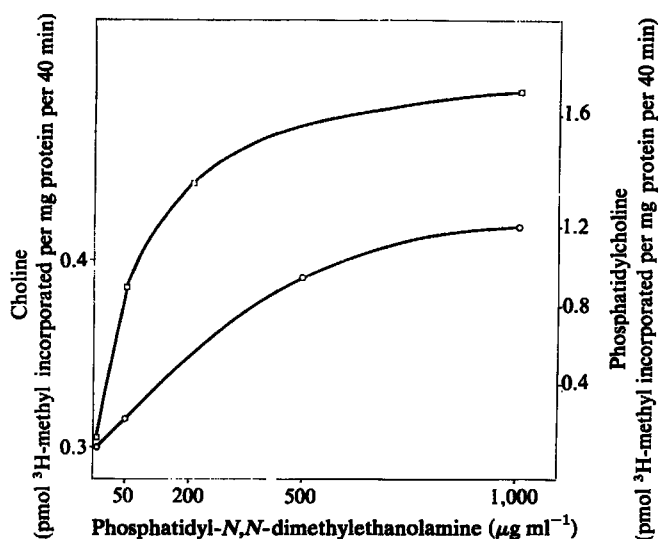


Fig. 2 Effects of P2 on PC (□) and choline (○) formation. Incubations were carried out for 40 min as described in Fig. 1 legend. Extractions and purification of products were performed as described in ref. 5 and Fig. 1 legend.

ATP-stimulated endoprotease is associated with the cell membrane of *E. coli*

Richard W. Voellmy* & Alfred L. Goldberg

Department of Physiology, Harvard Medical School, Boston, Massachusetts 02115, USA

Despite knowledge of the physiological significance and regulation of protein degradation in bacteria¹⁻⁷, the pathway of proteolysis and the responsible enzymes are still not known. Degradation of cell proteins in bacterial and animal cells requires continuous ATP production^{1,8}, inhibition of which in *Escherichia coli* prevents the degradation of normal proteins in growing cells⁸, accelerated breakdown of such proteins in starving cultures^{3,8,9} and the very rapid breakdown of abnormal proteins^{8,10}. Intracellular proteolysis proceeds by repeated endoproteolytic steps and ATP is required for the initial cleavages of the substrate¹⁰. We have recently demonstrated ATP stimulation of proteolysis in extracts of bacterial and animal cells¹¹⁻¹⁵. These ATP-stimulated systems seem to be responsible for the rapid degradation of abnormal proteins¹¹⁻¹⁴ *in vivo*, but they may also be involved in the catabolism of normal cell proteins^{1-3,9}, limited proteolysis, such as the processing of precursors for secreted or membrane proteins¹⁶⁻²⁰, and the selective inactivation of specific proteins, as occurs in the ATP-dependent cleavage of the lambda repressor by the recA protein²¹. We report here that membrane fragments contain an ATP-stimulated protease that degrades cell proteins to large peptides (of molecular weight (MW) 71,500) which are then rapidly hydrolysed to amino acids by soluble ATP-independent enzymes.

When extracts of freshly grown *E. coli* were prepared by gentle methods (by grinding with alumina, by using the French press at 3,000 p.s.i., or by freezing and thawing followed by lysozyme treatment and osmotic shock¹³), the degradation of [$Me-^{14}C$]globin^{13,22} was stimulated two- to fourfold by addition of ATP (3 mM). With less gentle methods of lysis (the French press at 12,000–15,000 p.s.i., repeated freezing and thawing followed by lysozyme treatment, or sonication), the resulting extracts contained high levels of proteolytic activity but showed very little (<25%) stimulation by ATP¹³.

Approximately 70–80% of the proteolytic activity after gentle cell lysis was recovered on the pellet resulting from centrifugation at 160,000g for 60 min (Table 1). When these pellets were resuspended, they showed ATP-stimulated proteolysis. The proteolytic activity remaining in the supernatant was not affected by ATP (Table 1).

When these extracts were analysed by gel filtration on Sepharose 2B or 4B, most of the ATP-stimulated activity was eluted in the void volume ahead of 70S ribosomes (Fig. 1), and is thus associated with large particulate material. Some soluble proteolytic activity eluted with the bulk of cell protein of molecular weights less than that of β -galactosidase (MW 540,000). This latter activity (see below and refs 29, 39, 40) showed a much smaller stimulation by ATP than did the peak in the void volume, in which ATP increased proteolysis two- to fivefold. Identical elution patterns were obtained when gel filtration was repeated on the material in the void volume or was performed in the presence of 3 mM ATP. Therefore, ATP does not stimulate proteolysis by releasing the proteolytic activity from a bound to a free state.

This high-molecular-weight fraction contains polyribosomes and membrane fragments, as indicated by a high content of the ($Ca^{2+} + Mg^{2+}$)ATPase (Fig. 2) and many membrane vesicles

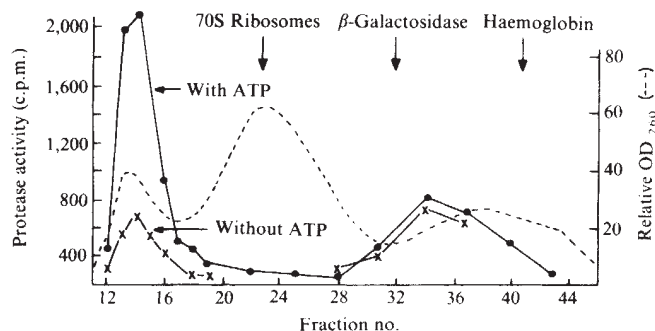


Fig. 1 Fractionation of *E. coli* proteolytic activity on Sepharose 4B. An extract from *E. coli*, prepared by grinding with alumina (50 mg protein), was filtered through an 80×2.5 cm column of Sepharose 4B equilibrated with 10 mM Tris-HCl, 10 mM $MgCl_2$, 50 mM KCl and 0.5 mM dithiothreitol (pH 7.5). Proteolytic activity was determined with (●) and without (×) an ATP-generating system as in Table 1 legend. β -galactosidase was assayed using *o*-nitrophenyl- β -D-galactoside as a substrate¹⁰. Similar data were obtained with extracts prepared as in Table 1 legend.

were evident in the electron microscope (in collaboration with Dr Richard Goldstein). When the extracts were incubated with puromycin ($50 \mu g ml^{-1}$) for 5–10 min at 37 °C or with ribonuclease A ($100 \mu g ml^{-1}$) before gel filtration to convert polysomes to monosomes²³, or when cells were incubated for 15 min at 10–15 °C before their collection to release ribosomes from the RNA²³, the ATP-stimulated proteolytic activity still eluted in the void volume of Sepharose 4B. Thus it could not have been associated with the polysomes.

Extracts subjected to sucrose density centrifugation showed maximal proteolytic activity at 47.5% sucrose (Fig. 2), the same density as membrane fragments prepared from marker cells grown in parallel in the presence of 3H -oleate to label membrane lipids. The maximal activity of the membrane-associated ($Ca^{2+} + Mg^{2+}$)ATPase was also found in these same lipid-containing fractions (Fig. 2).

If the ATP-stimulated protease is membrane-associated, it should co-precipitate on incubation with an antibody directed against another membrane protein. As membrane vesicles prepared by grinding or French press are inverted²⁴, an antibody against an inner membrane protein, the F_1 ATPase, was used. When it was incubated with the fraction purified by gel filtration (Fig. 1), it is precipitated most (76–90%) of the membrane-bound ATPase (data not shown) and reduced the proteolytic activity in the supernatant (Table 2). By contrast, when the antibody was added to the soluble proteolytic activity obtained by gel filtration (Fig. 1), no inhibition of globin hydrolysis was observed. A control immunoprecipitation was also carried out with purified exogenous elongation factor G (EFG) and an antibody directed against this protein (Table 2). The concentrations of EFG and anti-EFG were adjusted such that a similar amount of protein precipitated in these conditions to that which precipitated with the antibody against the membrane ATPase. Virtually none of the membrane-associated proteolytic activity was lost from the supernatant.

Previous studies^{1,25,26} have demonstrated appreciable proteolytic activity in membrane fractions prepared by low-speed centrifugation. We therefore tested whether this fraction also contained the ATP-stimulated proteolytic system. Cells were lysed by freezing and thawing followed by lysozyme treatment and osmotic shock, and the extracts were centrifuged at 3,000g for 10–15 min to remove unbroken cells and then at 25,000g for 20 min. The pellets of the second spin were very active in hydrolysing ^{14}C -globin, but this reaction was not stimulated by ATP (3 mM). However, in the presence of 10 mM sodium azide (an ATPase inhibitor), ATP caused a clear stimulation of protein degradation (data not shown), suggesting that a functioning ($Ca^{2+} + Mg^{2+}$)ATPase is not essential for the stimulation of proteolytic activity, in accord with findings on intact cells⁸.

* Present address: Department of Molecular Biology, University of Geneva, Geneva, Switzerland.

Table 1 Distribution of the ATP-stimulated proteolytic activity on ultracentrifugation

Fractions	Proteolytic activity (c.p.m.)		% Stimulation by ATP
	Without ATP	With ATP	
Crude extract	5,490	14,260	160
Pellet after second wash	3,831	8,772	129
Supernatant of first wash	1,850	1,860	—
Supernatant of second wash	804	792	—

E. coli K12 strain NF172 (ref. 9, 500-ml portions) was grown in Luria broth⁴² at 37 °C with vigorous aeration. The cells were cooled rapidly and collected at an A_{550} of 1.0. Three litres of cells were lysed by grinding with alumina for a short period. The resulting crude extract in buffer I (25 mM Tris-HCl pH 7.5, 20 mM $MgCl_2$, 0.5 mM dithiothreitol) was then centrifuged at 12,000g for 20 min. The supernatant was dialysed overnight against buffer I, and an aliquot of this crude extract was assayed for proteolytic activity. The dialysed extract was spun at 120,000g for 75 min. The pellets were resuspended in this buffer and centrifuged again. The pellet from this first wash was resuspended once more and centrifuged as before. The final pellets were resuspended in the original volume of buffer I, and together with the supernatants of the washes were assayed for proteolytic activity in the presence or absence of ATP and an ATP-generating system. The hydrolysis of [Me-¹⁴C]- (apo)haemoglobin^{13,22} to acid-soluble material was used to estimate proteolytic activity. Incubation mixtures contained 1 ml of extract, 10 μ l of ¹⁴C-globin (~15,000 c.p.m.) and 50 μ l of an ATP solution (60 mM) or an ATP-generating system (60 mM ATP, 250 mM phosphocreatine, 0.4 mg ml⁻¹ creatine phosphokinase, rabbit muscle type I [Sigma]). Incubations were carried out at 30 °C for 60–90 min. Comparisons between the absolute proteolytic activity in the crude extracts, supernatants and pellets may be misleading as these fractions contained very different amounts of nonradioactive proteins that may compete with the ¹⁴C-labelled substrate. Qualitatively, similar data were obtained when the pellets were resuspended in buffer I containing 0.5 M KCl or 1 mM Tris-HCl (pH 7.5), although these conditions reduced the degree of stimulation by ATP.

Furthermore, in related experiments (data not shown) we have demonstrated ATP-stimulated proteolysis in membrane vesicles of mutants²⁷ that lack a functioning ($Ca^{2+} + Mg^{2+}$) ATPase (AN120, an *uncA*⁻ strain).

The acid-soluble products of protein breakdown were analysed by gel filtration on Sephadex G-15. Crude *E. coli* extracts, prepared by alumina grinding, and the membrane fraction, isolated by gel filtration on Sepharose 4B, were incubated in parallel with either ¹⁴C-globin or ³⁵S-labelled *E. coli* protein (denatured with acid to inactivate proteases). Both substrates were hydrolysed on these extracts to acid-soluble material by an ATP-stimulated process (stimulation was 100–200% with globin and 70–100% with the *E. coli* proteins). The crude extracts degraded the ³⁵S-proteins to amino acids (Fig. 3) and thus can carry out the complete degradation of cell proteins, in similar fashion to intact cells. The [Me-¹⁴C]globin was hydrolysed to dipeptides. Related observations suggest that the labelled methyl group on the lysine prevents the hydrolysis of dipeptides by certain amino-peptidases.

When either substrate was incubated with the purified membrane fraction, most of the resulting acid-soluble ³⁵S- or ¹⁴C-products eluted close to the void volume. Very little was in the form of amino acids or even dipeptides (Fig. 3). The sizes of these proteolytic products did not differ whether or not ATP was present (data not shown). The ATP-stimulated proteolytic activity generated mainly peptides of MW 1,500 or greater²⁸ which were not degraded further. Thus the rate-limiting step in protein degradation in these extracts is an ATP-stimulated endoproteolytic cleavage, in accord with *in vivo* findings¹⁰. This initial step is membrane associated, while soluble cytoplasmic enzymes (exo- and endopeptidases^{6,29,30}) catalyse the subsequent conversion of the peptides to free amino acids. Thus, when the ³⁵S-labelled peptides generated by the membrane fraction were mixed with the soluble fraction obtained by

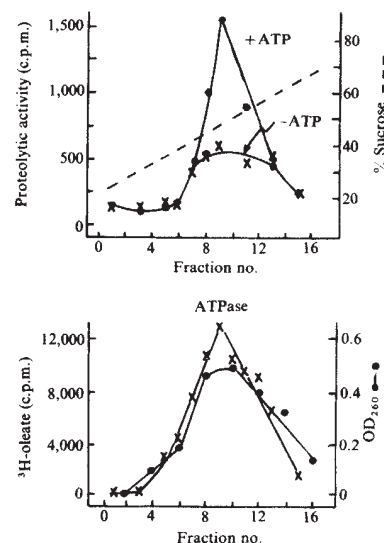


Fig. 2 Sucrose-density centrifugation of the particulate proteolytic activity. An extract prepared with the French press at 3,000 p.s.i. was centrifuged at 12,000g for 20 min. The supernatant was spun for 60 min at 12,000g and the resulting pellet washed once and resuspended in buffer I. For measurements of proteolytic activity against ¹⁴C-globin (see Table 1) and ATPase activity, 5–10 mg of the pelleted material were applied to linear sucrose gradients (32.5 ml) ranging from 25–65% and spun for 24 h at 25,000 r.p.m. in a SW27 rotor. Fractions (65 drops) dialysed against buffer I to remove the sucrose were assayed for proteolytic activity. To identify the fractions containing membrane lipids, cells were grown in medium containing ³H-oleate (28.5 mCi per mg, NEN). Extracts from the ³H-oleate-labelled cells were spun in parallel, and the radioactivity in each fraction measured (lower panel). To assay ATPase activity, 0.5-ml portions of the fractions were incubated with 0.5 mM ATP for 45 min at 30 °C. The samples were then boiled, and their ATP content measured using the firefly luciferase assay^{8,9}.

centrifugation (Table 1) of the crude extract, they were quickly hydrolysed to free amino acids. This latter reaction was not stimulated by ATP.

Although a large fraction, perhaps most, of the ATP-stimulated proteolytic activity is associated with the cell membrane, we cannot conclude what per cent of the proteolytic activity is actually membrane bound *in vivo* or even in these cell-free extracts, because the soluble and membrane fractions contain very different amounts of peptidases (Fig. 3) and of nonradioactive proteins that serve as competing substrates. A localization of the ATP-stimulated proteolysis on the inner cell membrane seems likely as alumina grinding and the French Press at low pressures create primarily inverted membrane vesicles, which should be impermeable to ATP and ¹⁴C-globin. Furthermore, protein breakdown in intact cells is completely dependent on the continuous production of ATP^{1,3,8–10} and therefore requires contact with the intracellular nucleotide pools. However, these studies have not eliminated conclusively a possible association with the outer cell membrane³¹. As the existence of proteases free within a cell might be harmful, an association with the membrane could reduce the frequency at which critical enzymes are inappropriately hydrolysed. This localization may also affect the selectivity of protein breakdown. Bacteria, like animal cells, degrade soluble proteins with highly abnormal conformations especially rapidly^{1,2,4–7,10–12}. Such partially unfolded proteins probably have more hydrophobic regions exposed on their surface than do normal ones^{1,12} and, therefore, may bind differentially to the cell membrane. It is perhaps relevant that when bacteria make large amounts of abnormal proteins, these molecules associate in intracellular aggregates that lie adjacent to the surface membrane^{32,33}.

On the cell membrane, the ATP-stimulated endoprotease may also be important in the turnover of membrane components

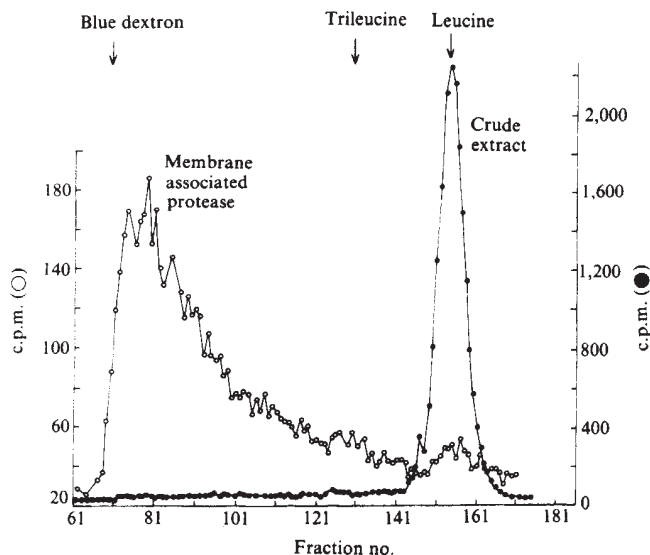


Fig. 3 Size distribution of the acid-soluble material produced by the membrane-associated protease and the crude extract. Cells were lysed by grinding with alumina. One portion of the extract was filtered through Sepharose 4B, and the fractions containing the membrane-associated protease were pooled and concentrated by ultrafiltration. This material (○) and a portion of the crude extract (●) containing an equal amount of protein were incubated with ^{35}S -labelled denatured *E. coli* protein for 2 h at 30 °C in the presence of an ATP-generating system. The protein was then precipitated with 10% trichloroacetic acid, and the supernatants applied to a 2 × 78 cm column of Sephadex G-15 in 0.05 M Tris-HCl (pH 7.5). Trileucine and leucine were run as molecular weight markers and were assayed with fluorescamine²⁶. To prepare ^{35}S -labelled *E. coli* proteins, 10 ml of strain NF 161 (*spoT*⁻, *arg*⁻, *met*⁻) were grown on glucose minimal medium⁸ containing arginine (60 $\mu\text{g ml}^{-1}$), a limiting amount of methionine (1.5 $\mu\text{g ml}^{-1}$) and 0.5 mCi of ^{35}S -methionine (460 Ci nmol⁻¹, NEN). When growth ceased, the cells were centrifuged and washed with 10 ml of growth medium lacking ^{35}S -methionine. The cells were then treated with 10% trichloroacetic acid. After centrifugation, the pellet was dissolved in 1 M NaOH and the pH adjusted to 7.0 with HCl. Before use, the suspension was dialysed extensively against buffer I.

and in the processing of secreted periplasmic or membrane proteins from larger precursors^{16-19,21}, a process believed to occur on the inner cell membrane^{17,21,37}. Inner membrane vesicles are able to process membrane protein^{17,20}, and this reaction seems to depend on the high-energy membrane state^{34,35}. Further characterization of the 'signal peptidase(s)' and the membrane-associated globin-degrading activities should clarify whether these proteolytic systems are in some way related. In addition, the ATP-stimulated enzyme may be involved in the rapid degradation of the released signal peptides.

These findings raise the possibility that the effect of ATP is mediated by some membrane-related function. However, there is strong evidence that the high-energy membrane state is not required for protein breakdown *in vivo*⁸, and uncoupling agents do not inhibit the ATP effect in these membrane preparations (data not shown). The present studies have also ruled out an involvement of the (Ca^{2+} + Mg^{2+})ATPase. Moreover, when *E. coli* are lysed by less gentle methods than were used here, we can isolate from the soluble fraction an ATP-dependent endoprotease, as well as seven ATP-independent proteases^{29,38-40}. Therefore, a membrane-association is not essential for the effect of ATP on protein degradation. After purification^{29,41}, the soluble ATP-dependent enzyme resembles strongly the membrane-associated activity described here²⁹. Possibly the soluble enzyme is derived from the membrane fraction, or alternatively the bacteria may contain two distinct ATP-stimulated proteolytic systems that serve different physiological functions.

Table 2 Loss of proteolytic activity on treatment with antibody against membrane-associated ATPase

Source of proteolytic activity	Additions	Activity remaining in supernatant (c.p.m.)	% Proteolytic activity lost from supernatant
Particulate fraction	Rabbit IgG	4,661	0
	anti-ATPase	2,876	39
	(EFG) + anti-EFG	4,455	4
Soluble fraction	Rabbit IgG	5,677	0
	Anti-ATPase	5,239	8

An extract prepared by grinding with alumina was chromatographed through Sepharose 4B. The peaks of proteolytic activity in the void volume (particulate activity) and in the soluble fraction were separated by gel chromatography as in Fig. 1. Portions (1 ml) of the two fractions in buffer I were incubated with antibody for 24 h at 4 °C. Precipitates were removed in a refrigerated clinical centrifuge, and proteolytic activity against ^{14}C -globin in the supernatants was assayed in the presence of an ATP-generating system (Table 1). The purified IgG fractions ($A_{280} = 0.57$) or purified EFG ($A_{280} = 0.02$) were added to 1 ml of the protease preparations. All antisera were purified by ammonium sulphate fractionation and chromatography on diethylaminoethyl and CM-cellulose⁴³ to remove the associated proteolytic activity. The purified IgG fractions contained no measurable activity against ^{14}C -globin.

We thank Dr P. C. Tai for advice and for providing pure ribosomes and ribosomal subunits, Dr F. Gibson for strain AN120, Professor L. Heppel for the antibody against membrane ATPase, and Ms Robin Levine for assisting in the preparation of this manuscript. During this work, R. W. V. was a Fellow of the Swiss National Fund. These studies were supported by grants from the National Institute of Neurobiological Diseases and Stroke and the Juvenile Diabetes Fund.

Received 15 May; accepted 12 November 1980.

- Goldberg, A. L. & St John, A. C. A. *Rev. Biochem.* **45**, 747-803 (1976).
- Pine, M. A. *Rev. Microbiol.* **26**, 103-126 (1972).
- Mandelstam, J. *Bact. Rev.* **24**, 289-308 (1960).
- Zipser, D. & Bhausaar, P. J. *Bact.* **127**, 1538-1542 (1976).
- Miller, C. G. in *Limited Proteolysis in Microorganisms* (eds Holtzer, H. & Cohen, G.) 65-69 (US Dept of Health, Education & Welfare, DHEW Publ. no. 79-1591, 1979).
- Olden, K. & Goldberg, A. L. *Biochem. biophys. Acta* **542**, 385-398 (1978).
- St John, A. C. & Goldberg, A. L. *J. biol. Chem.* **253**, 2705-2711 (1978).
- Kowit, J. D. & Goldberg, A. L. *J. biol. Chem.* **252**, 8350-8357 (1977).
- Etlinger, J. D. & Goldberg, A. L. *Proc. natn. Acad. Sci. U.S.A.* **74**, 54-58 (1977).
- Goldberg, A. L., Kowit, J. D. & Klemes, Y. in *Protein Turnover and Lysosomal Function* (eds Segal, H. & Doyle, D.) 171-196 (Academic, New York, 1978).
- Murakami, K., Voellmy, R. & Goldberg, A. L. *J. biol. Chem.* **254**, 8194-8200 (1979).
- Voellmy, R., Murakami, K. & Goldberg, A. L. in *Limited Proteolysis in Microorganisms* (eds Holtzer, H. & Cohen, G.) 7-17 (US Dept of Health, Education & Welfare, DHEW Publ. no. 79-1591, 1979).
- DeMartino, G. N. & Goldberg, A. L. *J. biol. Chem.* **254**, 3712 (1979).
- Blöbel, G. & Dobberstein, B. *J. Cell. Biol.* **67**, 852-862 (1975).
- Chang, C. N., Blöbel, G. & Model, P. *Proc. natn. Acad. Sci. U.S.A.* **75**, 361-365 (1978).
- Inouye, H. & Beckwith, J. *Proc. natn. Acad. Sci. U.S.A.* **74**, 1440-1444 (1977).
- Smith, W. P., Tai, P. L., Thompson, R. C. & Davis, B. D. *Proc. natn. Acad. Sci. U.S.A.* **74**, 2830-2834 (1977).
- Chang, C. N., Inouye, H., Model, P. & Beckwith, J. *J. Bact.* **142**, 726-728 (1980).
- Roberts, J. W., Roberts, C. W. & Mount, D. W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 2283-2287 (1977).
- Etlinger, J. D. & Goldberg, A. L. *J. biol. Chem.* **255**, 4563-4568 (1980).
- Subramanian, A. R., Davis, B. D. & Beller, R. J. in *Cold Spring Harb. Symp. quant. Biol.* **34**, 223-225 (1969).
- Futa, M. *J. Membrane Biol.* **15**, 15-28 (1974).
- Regnier, P. & Thang, M. N. *FEBS Lett.* **36**, 31-33 (1973).
- Kowit, J., Choy, W. N., Champe, S. P. & Goldberg, A. L. *J. Bact.* **128**, 776-784 (1976).
- Butlin, J. D., Cox, G. B. & Gibson, F. *Biochem. J.* **124**, 75-81 (1971).
- Reiland, J. *Meth. Enzym.* **22**, 287-321 (1971).
- Voellmy, R., Rosenthal, E., Gronastajski, R. & Goldberg, A. L. *Cell* (in the press).
- Miller, C. G. & Schwartz, G. J. *Bact.* **135**, 603-611 (1978).
- Osborn, M. J., Gander, J. E., Parisi, E. & Carson, T. J. *J. biol. Chem.* **247**, 3962-3972 (1972).
- Prouty, W. F. & Goldberg, A. L. *Nature new Biol.* **240**, 147-150 (1972).
- Prouty, W. F., Karnovsky, M. J. & Goldberg, A. L. *J. biol. Chem.* **250**, 1112-1122 (1975).
- Ito, K., Mandel, G. & Wickner, W. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1199-1203 (1979).
- Date, T., Zwizinski, C., Ludmerer, S. & Wickner, W. *Proc. natn. Acad. Sci. U.S.A.* **77**, 827-831 (1980).
- Nelson, N. & Schatz, G. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4365-4369 (1979).
- Zwizinski, C. & Wickner, W. *J. biol. Chem.* (in the press).
- Swamy, K. H. S. & Goldberg, A. L. (submitted).
- Goldberg, A. L., Swamy, K. H. S., Chung, C. H. & Larimore, F. S. in *Meth. Enzym.* (in the press).
- Goldberg, A. L., Strnad, N. P. & Swamy, K. H. S. *Ciba Fdn Symp.* **75**, 227-251 (1980).
- Larimore, F., Waxman, L. & Goldberg, A. L. (submitted).
- Miller, J. *Experiments in Molecular Genetics*, 433 (Cold Spring Harbor Laboratory, New York, 1972).
- Palacios, R., Palmer, R. D. & Schimke, R. T. *J. biol. Chem.* **247**, 2316-2321 (1972).

Two mutations that alter the regulatory activity of *E. coli* *recA* protein

Jeffrey W. Roberts & Christine W. Roberts

Section of Biochemistry, Molecular and Cell Biology,
Wing Hall, Cornell University, Ithaca, New York 14853, USA

The *Escherichia coli* *recA* gene product is both a direct participant and a central regulatory element in important processes of DNA repair, including one that is probably responsible for all radiation mutagenesis and some chemical mutagenesis. It has the direct function of catalysing pairing of single-stranded DNA to a homologous region in duplex DNA^{1,2}, a reaction thought to be fundamental to genetic recombination. This activity of *recA* protein probably contributes to DNA repair by promoting recombination between damaged DNA molecules³. We have shown that *recA* protein also has a regulatory function, mediated by its ability to destroy repressors and possibly other proteins by proteolytic cleavage, leading to the induction of certain genes⁴⁻⁶; these include *recA* itself, regulated by the *lexA* gene product⁷⁻⁹, other genes probably involved in DNA repair, some of which are also regulated by the *lexA* gene product¹⁰, and the early genes of temperate bacteriophages, which are regulated by their repressors. *recA* protein is activated to cleave repressors *in vitro* by interaction with ATP (or the ATP analogue ATP- γ S) and a polynucleotide such as single-stranded DNA⁴, probably the same interaction that initiates the strand-pairing reaction¹¹. We have proposed that *recA* protein is also activated to attack repressors *in vivo* when it interacts with single-stranded DNA; DNA-damaging treatments such as UV-ray irradiation would thereby invoke the *recA*-dependent functions of recombination, repair, mutagenesis and prophage induction⁴. As further evidence that the proteolytic activity of *recA* protein is responsible for its regulatory function, we show here that the ability of two mutationally altered *recA* proteins to cleave phage λ repressor correlates with the ability of the mutant cells to induce prophage.

Both mutations *recA142* and *recA430* (formerly named *lexB30*) prevent phage repressor inactivation after inducing treatments, but they affect *recA* function in the cell differently. (1) The mutant *recA142* is defective in genetic recombination, like classical *RecA*⁻ mutants, and it does not induce λ prophage after normal inducing treatments¹². Unlike most *RecA*⁻ mutations, however, *recA142* supports spontaneous induction, an unprovoked but *recA*-dependent basal level of induction that occurs in wild-type lysogens (ref. 12 and Table 1). A λ repressor fragment characteristic of *recA*-mediated cleavage results from this spontaneous induction in both wild-type and *recA142* lysogens¹³. We thus expected that *recA142* protein would be able to cleave λ repressor in some conditions *in vitro*, but might have some defect in the interaction with DNA thought to initiate both strand pairing and repressor cleavage. (2) The mutant *recA430* was selected by Devoret and Blanco^{14,15} to be defective in λ induction, and it neither responds to inducing agents nor displays spontaneous induction (ref. 14 and Table 1). Furthermore, *recA* protein is not amplified in a *recA430* mutant cell in response to inducing treatments^{16,17}. Although the *recA430* mutation maps in the *recA* gene^{17,18}, conjugation crosses in the *recA430* cell produce nearly normal numbers of recombinant progeny. We expected that *recA430* protein would be unable to cleave λ repressor, but might be normal in activities required for genetic recombination.

To provide a source of these mutant proteins, we used strains carrying the *lexA* gene mutation *spr-51* (ref. 7), which is thought to inactivate a repressor of the *recA* gene and thus cause overproduction of *recA* protein^{5,19}. We confirmed that these derivatives maintain the *recA* phenotypes of the original

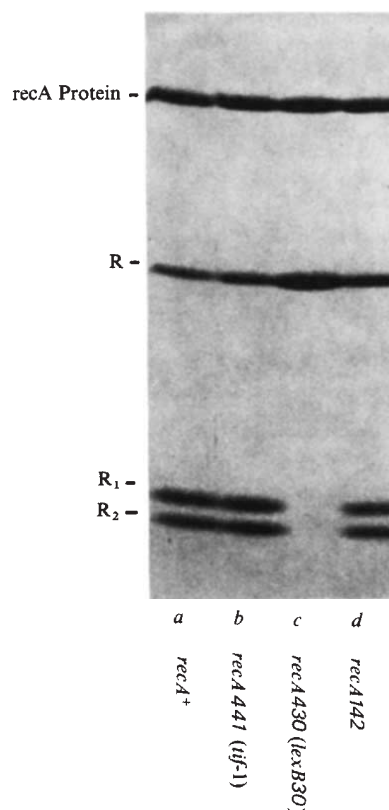


Fig. 1 Repressor cleavage directed by purified mutationally altered *recA* proteins. Reaction mixtures of 30 μ l were constructed as described previously⁴, except that they contained 2 mM MgCl₂, 1 mM ATP- γ S, 0.047 M NaCl, 0.013 M NaCl, 1.5 μ g *recA* protein, 3 μ g λ repressor and 0.039 μ g denatured λ DNA. After incubation for 3 h 20 min at 37 °C, the reaction mixtures were subjected to SDS-gel electrophoresis and the gel was stained as described previously⁴. The *recA* protein concentration was calculated from a measured absorbance at 280 nm of 0.60 units mg⁻¹ ml⁻¹, corrected from the previously assumed 1.0 units mg⁻¹ ml⁻¹ (ref. 4). Track a, *recA*⁺ protein; track b, *recA441* protein (formerly named *tif-1*); track c, *recA430* protein; track d, *recA142* protein. R, λ repressor; R₁ and R₂, proteolytic fragments of λ repressor. *recA* protein and λ repressor were prepared as described elsewhere⁴. *recA430* protein was prepared from the *spr-51 recA430* strain KM1193, obtained from K. McEntee. *recA142* protein was prepared from the *spr-51 recA142* strain JR235, constructed by P1 transduction of DM1420 (ref. 7) to *srl300::Tn10*, followed by transduction to *Sr1*⁺ by P1 grown on the *recA142* strain JC4728 and screening for sensitivity to UV-light irradiation.

mutants, in that a λ lysogen of *spr-51 recA142* displays spontaneous induction whereas a lysogen of *spr-51 recA430* does not (data not shown). Like its *spr*⁺*recA142* parent, the *spr-51 recA142* strain shows by patch tests the high sensitivity to killing by UV-ray irradiation characteristic of *recA*⁻ mutants (data not shown). The parental *spr*⁺*recA430* mutant is moderately UV sensitive¹⁵; its *spr-51* derivative is less sensitive by patch tests, but still distinctly more sensitive than wild type.

We tested the ability of *recA* protein purified from *recA142* and *recA430* bacteria to cleave λ repressor in the single-stranded DNA and ATP- γ S dependent reaction. Figure 1 shows that *recA142* protein is active, whereas *recA430* protein is not. In control reactions, both wild-type *recA*⁺ protein and protein from the *recA* mutant *tif-1* (refs 5, 20) (renamed *recA441*), which displays constitutive expression of *recA*-dependent functions, are active. From densitometric scanning of stained repressor fragments produced during a time course, we determined that the *recA142* protein cleaves λ repressor at ~75%

the rate of wild type in the presence of ATP- γ S (11 compared with 15 ng repressor cleaved per min per μ g *recA* protein). In contrast, Fig. 1 implies that *recA430* protein could cleave λ repressor at no more than 2% the rate of wild type. We tested the mutants in conditions that provide the highest cleavage rate with wild-type *recA* protein⁴, but observed no repressor cleavage by *recA430* protein in other conditions—ATP instead of ATP- γ S, higher concentrations of DNA, or oligonucleotides substituted for λ DNA.

recA protein is a single-stranded DNA-dependent ATPase^{4,21-23} whose activity is probably essential to strand exchange, as this requires ATP hydrolysis^{2,11}. Figure 2 shows that *recA430* protein has normal ATPase activity, consistent with its normal function in recombination. This result also shows that the protein we prepared was not destroyed during cell lysis and purification. The *recA430* protein also is equivalent to the wild type in DNA-dependent entrapment of radioactive ATP- γ S on a membrane filter (data not shown), and is active in reannealing homologous single strands of DNA, possibly a partial reaction of the strand-pairing activity⁹. Thus, we conclude that *recA430* protein is specifically defective in cleaving phage λ repressor, and that this defect accounts for the failure of λ to be induced in *recA430* bacteria.

The *recA142* protein, which does not effectively support recombination *in vivo*, is also an active ATPase *in vitro* (Fig. 2). However, its activity is much more sensitive to ionic strength than that of wild-type *recA* protein, displaying only 10% of the activity of the wild type at 0.20 M NaCl. This suggests a weaker interaction of the *recA142* protein with DNA in the presence of ATP that could account both for the recombination deficiency of the mutant and for its failure to cleave λ repressor in inducing conditions: it may be inefficient in forming the ternary complex with ATP and DNA postulated to be the active intermediate in both strand pairing and repressor cleavage. In support of this interpretation, we have found that ATP, in contrast to the analogue ATP- γ S, promotes λ repressor cleavage much less efficiently with *recA142* protein than with wild-type *recA*⁺ protein (<0.2 compared with 2 ng repressor cleaved per min per μ g *recA* protein). We presume that *recA142* protein shows substantial cleavage activity with ATP- γ S because this analogue, unlike ATP, interacts essentially irreversibly with *recA* protein (unpublished experiments) and thus overcomes the reluctance of *recA142* protein to form the active ternary complex.

The ATPase activity of the *recA430* protein is only slightly more sensitive to ionic strength than the wild type (Fig. 2), presumably corresponding to its nearly normal activity in genetic recombination.

These experiments provide further direct evidence that cleavage of repressors by *recA* protein is the mechanism of repressor inactivation in induction. Because the *recA430* protein supports recombination but not λ repressor inactivation *in vivo*, induction cannot result from the activity of *recA* protein as

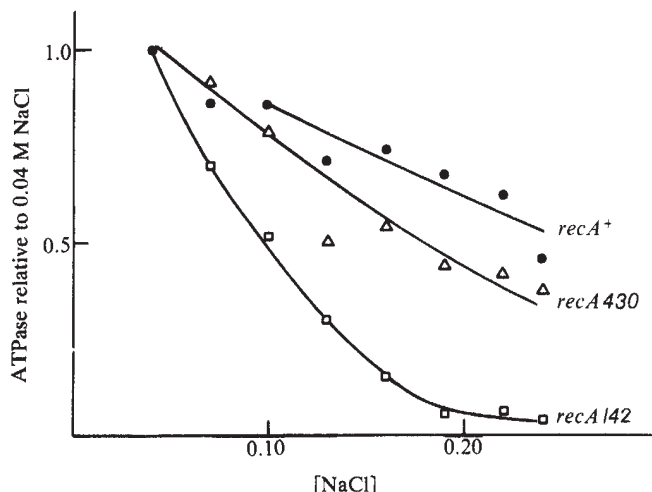


Fig. 2 Sensitivity of ATPase activity of mutationally altered *recA* proteins to ionic strength. The DNA-dependent ATPase activity was measured as described previously⁴, except that the concentration of *recA* protein was 100 μ g ml⁻¹ and the time of incubation was 30 min. Activities measured at 0.04 M NaCl were: *recA*⁺, 0.31 μ mol min⁻¹ mg⁻¹; *recA430*, 0.30 μ mol min⁻¹ mg⁻¹; *recA142*, 0.39 μ mol min⁻¹ mg⁻¹.

a recombination enzyme. Our results argue against models that postulate that the interaction of λ repressor with DNA (or a *recA*-DNA complex) is sufficient to promote inactivation of λ repressor in the absence of cleavage²⁴.

Besides their defect in λ repressor inactivation, *recA430* bacteria also fail to carry out other *recA*-dependent processes, such as amplification of *recA* protein and radiation mutagenesis^{17,25}. We suggest that these defects arise at least in part from the inability of *recA430* protein to inactivate the *E. coli* *lexA* protein, which is thought to be a repressor of the *recA* gene. The *lexA* protein is cleaved by *recA* protein⁸ in the conditions we have described for λ repressor inactivation; this supports the model that *recA* protein amplification occurs when *lexA* protein is cleaved in inducing conditions by *recA* protein present at a low basal level. Failure of *recA430* protein to cleave *lexA* protein would thus prevent amplification of *recA* protein. The inability of the *recA430* cell to be mutagenized by UV light could result from its failure to induce genes required for UV mutagenesis that are repressed by *lexA* or are regulated by other proteins sensitive to cleavage by *recA* protein. This is supported by the discovery of Kenyon and Walker¹⁰ that numerous *E. coli* genes involved in DNA repair are repressed by *lexA* and require *recA* function to be induced; these might include genes required for UV mutagenesis. Although the possibility has not been tested, it is likely that a third *recA*-dependent activity, lethal filamentation, is defective in *recA430*, as this process is also thought to result from inactivation of *lexA* protein by *recA* protein^{9,26}.

We thank N. L. Craig and E. M. Phizicky for advice and discussions, N. L. Craig for performing some preliminary experiments, A. J. Clark, R. Devoret, D. Botstein, D. Mount and K. McEntee for bacterial strains, and V. Vogt for criticizing the manuscript. This work was supported by NIH grant GM21941 and by a NIH Career Development Award to J.W.R.

Received 7 October 1980; accepted 29 January 1981.

- Shibata, T., DasGupta, C., Cunningham, R. & Radding, C. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1638-1642 (1979).
- McEntee, K., Weinstock, G. & Lehman, I. *Proc. natn. Acad. Sci. U.S.A.* **76**, 2615-2519 (1979).
- Rupp, W. & Howard-Flanders, P. *J. molec. Biol.* **31**, 291-304 (1968).
- Craig, N. L. & Roberts, J. W. *Nature* **283**, 26-30 (1980).
- Roberts, J., Roberts, C. & Craig, N. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4714-4718 (1978).
- Phizicky, E. M. & Roberts, J. W. *J. molec. Biol.* **139**, 319-328 (1980).
- Mount, D. *Proc. natn. Acad. Sci. U.S.A.* **74**, 300-304 (1977).

Table 1 Spontaneous and mitomycin C induction of λ in *recA* lysogens

Strain	No. of infectious centres per ml of culture	
	Spontaneous	Mitomycin C treated
<i>recA</i> ⁺ (λ ⁺)	1 $\times$ 10 ⁶	1 $\times$ 10 ⁹
<i>recA12</i> (λ ⁺)	<10 ¹	<10 ¹
<i>recA142</i> (λ ⁺)	5 $\times$ 10 ⁵	7 $\times$ 10 ⁴
<i>recA430</i> (λ ⁺)	2 $\times$ 10 ² (clear)	2 $\times$ 10 ² (clear)

Cultures were grown in minimal medium¹³ at 37 °C to 5 $\times$ 10⁷ cells per ml. Portions of each were grown a further 90 min in the presence and absence of 5 μ g ml⁻¹ mitomycin C, treated with chloroform and titred. Phage from the *recA430*(λ ⁺) culture made clear plaques, indicating that they were derived not from induction but from growth following mutational alteration of the repressor. Bacterial strains have been described elsewhere¹³, except the *recA430* strain JM776, which was obtained from R. Devoret. *recA12* is a standard *RecA*⁻ mutation.

8. Little, J. et al. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3225–3229 (1980).
9. Brent, R. & Ptashne, M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1932–1936 (1980).
10. Kenyon, C. & Walker, G. C. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2918–2923 (1980).
11. Cunningham, R. P., Shibata, T., DasGupta, C. & Radding, C. M. *Nature* **281**, 191–195 (1979).
12. Clark, A. J. A. *Rev. Genet.* **7**, 67–86 (1973).
13. Roberts, J. W. & Roberts, C. W. *Proc. natn. Acad. Sci. U.S.A.* **72**, 147–151 (1975).
14. Devoret, R. & Blanco, M. *Molec. gen. Genet.* **107**, 272–280 (1970).
15. Morand, P., Blanco, M. & Devoret, R. *J. Bact.* **131**, 572–582 (1977).
16. West, S. C. & Emmerson, P. *Molec. gen. Genet.* **151**, 51–67 (1977).
17. McEntee, K. *ICN-UCLA Symp.* **9**, 349–360 (1978).
18. Morand, P., Goze, A. & Devoret, R. *Molec. gen. Genet.* **157**, 69–82 (1977).
19. Gudas, L. & Mount, D. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5280–5284 (1977).
20. Castellazzi, M., George, J. & Buttin, G. *Molec. gen. Genet.* **119**, 139–152 (1972).
21. Roberts, J., Roberts, C., Craig, N. & Phizicky, E. *Cold Spring Harb. Symp. quant. Biol.* **43**, 917–920 (1978).
22. Ogawa, T. et al. *Cold Spring Harb. Symp. quant. Biol.* **43**, 909–916 (1978).
23. Weinstock, G., McEntee, K. & Lehman, I. *Proc. natn. Acad. Sci. U.S.A.* **76**, 126–130 (1979).
24. Sussman, R., Resnick, J., Calamne, K. & Baluch, J. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5817–5821 (1978).
25. Walsh, N. & Stocker, B. A. D. *J. Bact.* **137**, 830–838 (1979).
26. Mount, D. W., Walker, A. C. & Kosel, C. L. *J. Bact.* **116**, 950–956 (1973).

Nitrogen fixation gene (*nifL*) involved in oxygen regulation of nitrogenase synthesis in *K. pneumoniae*

Susan Hill, Christina Kennedy & Eugene Kavanagh

ARC Unit of Nitrogen Fixation, University of Sussex, Brighton BN1 9RQ, UK

Richard B. Goldberg & Robert Hanau

Department of Biology, Temple University, Philadelphia, Pennsylvania 19122, USA

The enzyme complex nitrogenase, which reduces N_2 to NH_4^+ , involves two redox proteins, both irreversibly damaged by O_2 (ref. 1). Enzyme activity therefore requires anaerobic conditions, a source of reductant and a large amount of ATP (~16 ATPs per N_2)^{2,3}. In both aerobic and facultative anaerobic N_2 -fixing bacteria, nitrogenase synthesis is regulated by O_2 and NH_4^+ , but in the aerobes there are also processes to protect the enzyme from O_2 damage^{4,5}. The mechanisms of repression by O_2 and NH_4^+ seem to be independent in the organisms so far examined^{6–8}. In the facultative anaerobe, *Klebsiella pneumoniae*, O_2 was shown to repress nitrogenase synthesis in an NH_4^+ -constitutive strain⁶. The fusion of the *Escherichia coli lacZ* gene into each transcriptional unit of the nitrogen fixation (*nif*) gene cluster in *K. pneumoniae* has facilitated studies with O_2 , because expression from the various *nif* promoters results in an O_2 -stable product (β -galactosidase). Notably, the *nifHDK* operon (the nitrogenase structural genes) was more sensitive to O_2 repression than the *nifLA* operon (regulatory genes)⁹. The characterization of mutants, reported here, indicates the involvement of a *nif*-regulatory gene product in the mechanism of O_2 control of nitrogenase synthesis.

The *nif-lac* fusions constructed by Dixon *et al.*⁹ with phage Mud(Aplac), in which *lac* is expressed from *nif* promoters, could not be directly used for the isolation of regulatory mutants, as this phage can re-transpose to new locations at high frequency (about 10^{-4} per generation; R. Dixon, personal communication). Therefore our approach was guided by the following argument. All previously described *nifL* mutations (except *nifL2265*) have been shown to be polar on the *nifA* gene¹⁰, the product of which is required for expression of all other *nif* transcriptional units⁹. *NifL2265* showed a partial *trans* dominance to pRD1 (ref. 11) (a plasmid carrying the entire wild-type *nif* cluster) which is consistent with a modified *nifL* gene product acting as a negative effector. Also, reversion to Nif^+ by strains carrying Mu insertions in *nifL* probably occurs by the imprecise excision of Mu, implying that *nifL* product is not necessary for nitrogenase synthesis¹⁰. Hence revertants of *nifL*

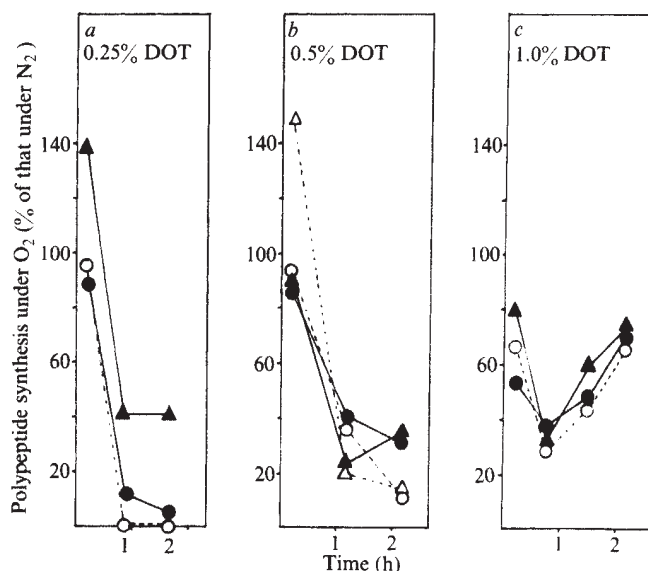


Fig. 1 Influence of DOT on synthesis of *nifH* (▲), *nifD* (●), *nifK* (○) and *nifJ* (△) gene products in the wild-type UNF50231 (a, b) and the mutant CK400 (c) strains. Results show the extent of *nif* gene product synthesis in organisms exposed to O_2 , expressed as a percentage of that occurring in a parallel culture in N_2 . Synthesis of *nif* polypeptides was quantified as the proportion of total polypeptide synthesized, and in cultures in N_2 the initial proportions for each *nif* product ranged as follows: *nifH*, 0.07–0.19; *nifD*, 0.05–0.13; *nifK*, 0.05–0.10; and *nifJ*, 0.02. Total protein synthesis was not markedly affected by exposure to O_2 . Similar experiments were performed on strains UN4357 and UNF753 and *nif* polypeptide synthesis after about 2 h exposure to 0.5% DOT compared with that occurring in N_2 was respectively as follows: *nifH*, 85 and 76%; *nifD*, 96 and 72%; *nifK*, 30 and 36%; and *nifJ*, 180 and 78% (similar results were obtained for strain UN4374); these values are significantly higher than the ones for the wild type shown in b of 35, 33, 12 and 13% for *nifH*, *nifD*, *nifK* and *nifJ* polypeptides respectively. In all experiments cultures were derepressed in an NH_4^+ -free medium²² in N_2 at 29 °C. A 20-ml sample was transferred anaerobically to the N_2 -flushed oxystat vessel, equipped with a calibrated O_2 electrode mounted in a side arm and a magnetic stirrer. Immediately after transfer air was slowly introduced into the N_2 supply to the vessel until the desired DOT was reached (never taking longer than 12 min), when the oxystat (Western Biological) was switched on. The DOT value, maintained ($\pm 20\%$) by control of the gas mixture, occasionally drifted, but the time during which the DOT fell below 0.25% (the lowest DOT tested and found to repress nitrogenase synthesis) never exceeded 8% of the total time of an experiment. A control 20-ml sample was stirred in N_2 in a vessel similar to and alongside the oxystat vessel in a water bath maintained at 30 °C. Samples (1 ml), removed at time intervals, were labelled with ^{14}C amino acids and protein subsequently analysed and quantified by SDS-polyacrylamide gel electrophoresis and densitometry of autoradiograms as described previously⁸.

mutants might yield strains in which the *nifA* gene product is synthesized but *nifL* gene product is inactive or lacking. Thus we screened Nif^+ revertants of strains carrying mutations in *nifL* for the failure of O_2 to repress synthesis of nitrogenase polypeptides.

Initially we investigated the influence of dissolved O_2 tension (DOT) on nitrogenase synthesis in wild-type strain UNF 50231. As expected, a DOT of 0.5%, which inhibited expression of β -galactosidase in strains carrying a *nifH::lac* fusion, repressed synthesis of nitrogenase polypeptides, as did a DOT of 0.25% (Fig. 1a, b). Besides the products of the *nifH*, *nifD* and *nifK* genes, the rate of synthesis of the *nifJ* gene product also declined (Fig. 1b).

A Nif^+ revertant (strain CK400) derived from a *nifL* point mutant (*nifL2265*) showed some escape from repression when exposed to 0.5% DOT. The kinetics of nitrogenase polypeptide synthesis in 1% DOT (Fig. 1c) confirmed that O_2 repression

was less severe than in the wild-type strain UNF50231 (Fig. 1a, b). Two other *Nif*⁺ revertants (UN4357 and UN4374; W. J. Brill, unpublished), derived from two strains carrying *Mu* insertions in *nifL*, behaved similarly (see Fig. 1 legend) and synthesis of the *nifJ* product also escaped repression in all three *Nif*⁺ revertants (see Fig. 1 legend).

Both the *Nif*⁺ and O₂ constitutive properties of the revertant CK400 were due to a second mutation (*nif2400*) mapping within the *nif* region; a His⁺*Nif*⁺ transductant (UNF753) of a (*his-nifM*) deletion mutant infected with phage P1 propagated in CK400 showed escape from O₂ repression (see Fig. 1 legend). To facilitate further physiological and genetic experiments, the *his*-linked mutations in CK400 were transduced into both *nifH::lac* and *nifN::lac* fusion strains. As expected, for both fusion alleles β -galactosidase activity could be detected by blue colour formed by transductants growing on NH₄⁺-free X-gal plates incubated in air (see Fig. 2 legend), and derepression of

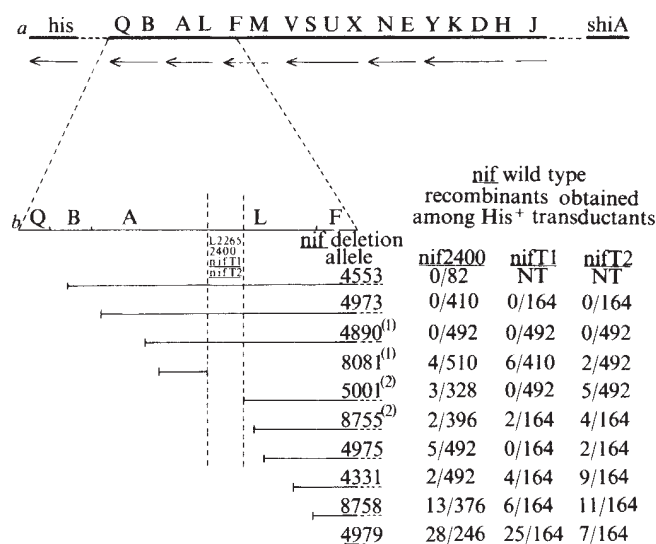


Fig. 2 Genetic mapping of O₂ constitutive *nif* mutations. *a*, Current map of the *nif* genes^{10,23,24}. The lines beneath the map indicate probable transcriptional units and the arrows show the direction of transcription. *b*, The O₂ constitutive *nifH::lac* strains described in the text were made His⁺ and used as recipients in phage P1Cmcl100- or P1Cmcl100-mediated transductions from the set of His⁺ *nif* deletion mutants^{10,24} shown above. His⁺ transductants were selected and scored for *lac* expression on solidified NH₄⁺-free medium²² agar plates containing 30 μ g X-gal (5-bromo-4-chloro- β -D-galactoside) and 50 μ g serine per ml incubated at 29°C in air; the mutants appeared blue whereas wild-type *nifH::lac* recombinants were colourless on aerobic plates. Possible wild-type transductants were always checked for retention of *nifH::lac* on anaerobic X-gal plates where they turn blue. If wild-type recombinants appeared among transductants, then wild-type DNA corresponding to the position of mutations responsible for O₂ constitutivity was not deleted from the donor deletion strains. If no such transductants appeared, then the mutations lie within the deleted *nif* segment. The number of transductants needed to be scored for *lac* phenotype (~400) to detect any wild-type recombinants was indicated from a similar deletion mapping experiment with a strain carrying the *nifL2265* allele in which 3.5 and 0.5% of His⁺ transductants were Nif⁺ using the *nif8755* and *nif8081* deletion alleles, respectively. In the pairs of deletion mutations marked (1) and (2) the relationship between the *his*-proximal ends is uncertain. The *his*-distal end of *nif8081* is near the *nifAL* intergenic region and probably within *nifL* as no *nifA* mutations have been found to map in the deletion interval defined by the ends of the *nif8081* and *nif8755* mutations. The placement of *nifL2265* within this deletion interval conflicts with other reported mapping experiments^{10,24}; however, repeated experiments where His⁺ transductants were scored for Nif, rather than being directly selected on NH₄⁺-free medium which is less efficient, consistently gave this result. The *nifT*⁻ mutations T1 and T2 were derived from strains Kp5161-3 and Kp5160-3, respectively^{13,25}. NT, not tested.

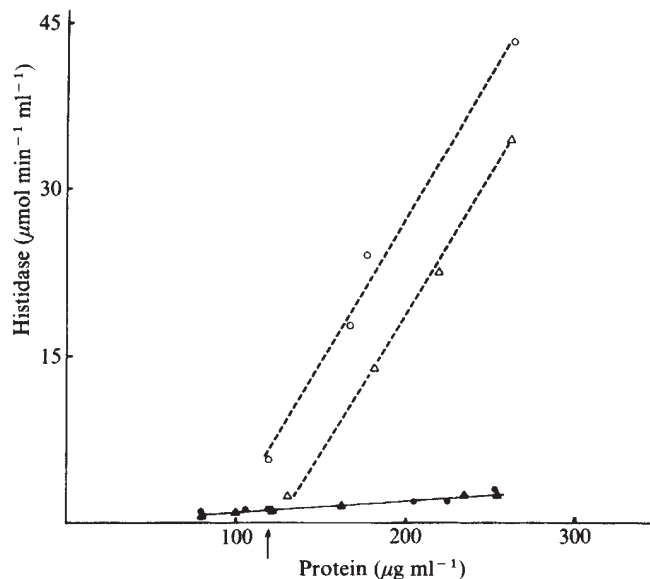


Fig. 3 Influence of an anaerobic to aerobic shift on the expression of histidase in the wild-type strain KG7035 (triangles) and the mutant UNF753, described in the text (circles). Cultures were grown in anaerobic conditions (closed symbols), and when the total protein concentration had reached 120 μ g ml⁻¹, indicated by the arrow, the culture was divided and half was transferred to an aerobic environment (open symbols)¹⁹.

β -galactosidase also occurred in liquid NH₄⁺-limited cultures either maintained in a DOT of 0.5% or bubbled with air (Table 1). In contrast, the parent fusion strains were colourless on aerobic X-gal plates and failed to derepress β -galactosidase in 0.5% DOT or air bubbling (Table 1). The apparent insensitivity of β -galactosidase derepression to O₂ in the *nifH::lac* fusion strain carrying the *his*-linked mutations in CK400 was not due to an altered respiratory activity; the potential respiratory rate¹² and the O₂ demand in a DOT of 0.5% were the same in the mutant and parent strains (data not shown).

Three other suspected *nif* regulatory mutations including a previously mentioned Nif⁺ revertant of a *nifL::Mu* strain (UN4357) and two mutants designated '*nifT*⁻'¹³ also allowed synthesis of β -galactosidase in air when linked to *nifH::lac* or to *nifN::lac* (Table 1). Significantly NH₄⁺ repressed *nifH::lac* and *nifN::lac* in all four mutants, and O₂ repression was restored in each *nifH::lac* strain by the introduction of the plasmid pRD1 (Table 1).

The blue phenotype on NH₄⁺-free X-gal indicator plates incubated in air was used to locate the *nif2400* and the two *nifT*⁻ mutations, and deletion mapping experiments (Fig. 2) indicate that all three mutations lie within the *nifLA* operon and probably within the *nifL* gene. The *nifT*⁻ mutants were isolated as Nif⁺ colonies from a glutamine synthetase mutant with impaired regulation for *nif* and *hut* (histidine utilization)¹³. Hence the *nifLA* operon seems to be involved in the control of *nif* expression by both O₂ and, as suggested previously⁹, the regulatory proteins of nitrogen assimilation¹⁴.

Several enzyme systems in the Enterobacteriaceae are apparently regulated by O₂ (refs 14–18) and in *K. pneumoniae*, *hut*¹⁹ and *nif*⁸ are controlled by both O₂ and NH₄⁺. We therefore examined the effect of the *his*-linked mutations in CK400 on *hut* expression; in both wild-type and mutant strains, O₂ was needed for histidase derepression (Fig. 3). Furthermore, these mutations did not alter the transient influence of O₂ on glutamine synthetase levels¹⁹ (data not shown).

Thus five different strains had *nif* mutations that relieved O₂ repression of the *nifHDK* and, where tested, the *nifEN* and the *nifJ* transcriptional units. In at least one of the strains (CK400) O₂ regulation of another operon outside *nif* was unaffected. The mutations mapped in the same deletion interval that includes the *his*-proximal end of the *nifL* gene. The *trans* dominance of

Table 1 Influence of O₂ on β -galactosidase derepression in strains carrying *nif-lac* fusions

Strain no.	Relevant chromosomal marker in addition to <i>recA56</i> and <i>lac</i> fusion allele <i>nifH2783</i> or <i>nifN2790</i>	Plasmid	Units of β -galactosidase activity after 6 h in		
			0.5% DOT	N ₂	N ₂ plus NH ₄ ⁺ (19 mM)
With <i>nifH2783</i>					
UNF762	None	None	12 ± 9	120 ± 2	1 ± 1
UNF777	<i>nifL2265</i> , <i>nif2400</i> (mutations from strain CK400)	None	156 ± 13	110 ± 30	5 ± 2
UNF767	<i>hisD</i>	pRD1	16	889	5
UNF798	<i>hisD</i> , <i>nifL2265</i> , <i>nif2400</i>	pRD1	27	864	2
			Air bubbling	N ₂ bubbling	N ₂ bubbling plus NH ₄ ⁺ (23 mM)
UNF762	None	None	9 ± 0	129 ± 17	ND
UNF777	<i>nifL2265</i> , <i>nif2400</i>	None	52 ± 3	71 ± 3	ND
UNF954	<i>hisD</i> (<i>nif</i> mutations from strain UN4357)*	None	335 ± 55	191 ± 17	2
UNF954	<i>hisD</i> (<i>nif</i> mutations from strain UN4357)*	pRD1	15 ± 2	622 ± 36	2
UNF951	<i>hisD nifT1</i> ⁻	None	129 ± 4	231 ± 10	2
UNF951	<i>hisD nifT1</i> ⁻	pRD1	5 ± 0	452 ± 52	2
UNF952	<i>hisD nifT2</i> ⁻	None	100 ± 12	261 ± 8	3
UNF952	<i>hisD nifT2</i> ⁻	pRD1	9 ± 0	366 ± 8	3
With <i>nifN2790</i>					
UNF979	None	None	6 ± 0	51 ± 1	2
UNF982	<i>nifL2265</i> , <i>nif2400</i>	None	90 ± 10	99 ± 7	3
UNF984	<i>hisD</i> (<i>nif</i> mutations from strain UN4357)*	None	68 ± 8	88 ± 12	3
UNF986	<i>hisD nifT1</i> ⁻	None	80 ± 8	87 ± 3	2
UNF988	<i>hisD nifT2</i> ⁻	None	33 ± 3	124 ± 9	3

NH₄⁺-grown cultures were suspended in an NH₄⁺-free medium and β -galactosidase derepression was followed as described previously⁹. Cultures were maintained in 0.5% DOT as described in Fig. 1 legend. ND, not determined.

* Strain UN4357 is a Nif⁺ derivative of strain UN2464 carrying *nifL5037::Mu*¹⁰.

wild-type *nif* and the relative insensitivity to O₂ of *nifL* expression⁹ suggest that the *nifL* product is involved as a negative effector in the O₂-regulatory mechanism specific to nitrogen fixation. Conversion to a negative effector may involve processes associated with O₂ metabolism rather than the direct interaction with the O₂ molecule. The mechanism of action of such a negative effector could be: (1) to prevent *nifA* expression; (2) to inactivate the *nifA* product; (3) to bind at each *nif* operator/promoter; (4) to promote *nif* mRNA breakdown; (5) to prevent *nif* translation; or (6) to promote *nif* gene product breakdown. The first seems unlikely because in strains carrying a *nifA::lac* fusion (where *nifL* is expressed) β -galactosidase derepression was not inhibited in 0.5% DOT (unpublished results). The sixth is improbable as synthesis of a hybrid polypeptide is unlikely in strains carrying the *nifH::lac* or the *nifN::lac* transcriptional fusions^{9,20,21}.

This work was supported in part by NSF grants PCM76-01915 and PCM79-22122 to R.B.G. We thank Dr S. Kumar for help in *nifL2265* mapping experiments, Drs F. Ausubel, C. Elmerich, W. Brill and D. MacNeil for strains, Professor J. R. Postgate and Drs M. Merrick and B. E. Smith for reading the manuscript, and L. Batsleer for technical assistance.

Received 18 August 1980; accepted 21 January 1981.

- Eady, R. R. & Postgate, J. R. *Nature* **249**, 805-810 (1974).
- Ljones, T. *FEBS Lett.* **98**, 1-8 (1979).
- Andersen, K., Shanmugam, K. T. & Valentine, R. C. in *Genetic Engineering for Nitrogen Fixation* (ed. Hollaender, A.) 95-110 (Plenum, New York, 1977).
- Kennedy, C. & Eady, R. R. in *Nitrogen Assimilation of Plants* (eds Hewitt, E. J. & Cutting, C. V.) 73-84 (Academic, London, 1979).
- Robson, R. L. & Postgate, J. R. *Rev. Microbiol.* **34**, 183-207 (1980).
- Postgate, J. R. *et al.* in *Biological Metabolism of Inorganic Nitrogen and Sulfur Compounds* (eds Bothe, H. & Trebst, A.) 103-115 (Springer, Heidelberg, in the press).
- Bergersen, F. J., Turner, G. L., Gibson, A. H. & Dudman, W. F. *Biochim. biophys. Acta* **444**, 164-174 (1976).
- Eady, R. R., Issack, R., Kennedy, C., Postgate, J. R. & Ratcliffe, H. J. *gen. Microbiol.* **104**, 277-285 (1978).
- Dixon, R. *et al.* *Nature* **286**, 128-132 (1980).
- MacNeil, T., MacNeil, D., Roberts, G. P., Supiano, M. A. & Brill, W. J. *J. Bact.* **136**, 253-266 (1978).
- Kennedy, C. *Molec. gen. Genet.* **157**, 199-204 (1977).
- Hill, S. J. *gen. Microbiol.* **93**, 335-345 (1976).
- Ausubel, F., Riedel, G., Cannon, F., Peskin, A. & Margolskee, R. in *Genetic Engineering for Nitrogen Fixation* (ed. Hollaender, A.) 111-128 (Plenum, New York, 1977).
- Leonardo, J. M. & Goldberg, R. B. *J. Bact.* **142**, 99-110 (1980).
- Pichinoty, F. *Biochim. biophys. Acta* **64**, 111-119 (1962).
- Spencer, M. E. & Guest, J. R. *J. Bact.* **114**, 563-570 (1973).
- Harrison, D. E. *Adv. microb. Physiol.* **14**, 243-314 (1976).
- Fimmel, A. L. & Haddock, B. A. *J. Bact.* **138**, 726-730 (1979).
- Goldberg, R. B. & Hanau, R. J. *Bact.* **141**, 745-750 (1980).

- Cassadaban, M. J. & Cohen, S. N. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4530-4533 (1979).
- Cassadaban, M. J., Chou, J. & Cohen, S. N. *J. Bact.* **143**, 971-980 (1980).
- Cannon, F. C., Dixon, R. A., Postgate, J. R. & Primrose, S. B. *J. gen. Microbiol.* **80**, 227-239 (1974).
- Puhler, A. & Klipp, W. in *Biological Metabolism of Inorganic Nitrogen and Sulfur Compounds* (eds Bothe, H. & Trebst, A.) (Springer, Heidelberg, in the press).
- Merrick, M. *et al.* *J. gen. Microbiol.* **117**, 509-520 (1980).
- Ausubel, F. M., Margolskee, R. F. & Maizels, N. in *Recent Developments in Nitrogen Fixation* (eds Newton, W., Postgate, J. R. & Rodriguez-Barrueco, C.) 347-356 (Academic, London, 1977).

Erratum

In the article 'Generation of F₁ hybrid cytotoxic T lymphocytes specific for self H-2' by K. Nakano *et al.*, *Nature* **289**, 559-563 (1981), Table 1 contained some errors. The correct version appears below.

Table 1 Genetics of target determinants for the lytic activity of B6D2F₁ anti-B6 CTLs

Group no.	Strain	Target cells		% Specific lysis (effector: target cells = 40:1)
		H-2 haplotype	K I SD AB J EC	
1	B6	b b b b b b b b		36.0
	B10	b b b b b b b b		32.4
	C3H.B10	b b b b b b b b		34.3
2	A	k k k k k d d d		9.0
	B10.A	k k k k k d d d		8.2
	DBA/2	d d d d d d d d		5.8
	B10.D2	d d d d d d d d		4.4
	C3H	k k k k k k k k		12.7
3	B10.A(4R)	k k b b b b b b		23.2
	B10.A(2R)	k k k k k d d d		26.0
	HTG	d d d d d d d d		38.0
	B10.A(5R)	b b b k k d d d		8.4
4	B6 × DBA/2	b b b b b b b b		8.6
		d d d d d d d d		
	B6 × C3H	b b b b b b b b		11.1
5		k k k k k k k k		
	B10 × B10.A(2R)	b b b b b b b b		31.1
		k k k k k d d d		
	B10.A × B10.A(2R)	k k k k k d d d		14.1
		k k k k k d d d		
	B10.A × B10.A(5R)	k k k k k d d d		13.4
		b b b k k d d d		

MATTERS ARISING

'Lipidic particles' are intermembrane attachment sites

THE nature of lipidic 'particles' and 'pits' in freeze-fracture replicas of liposomes has been much debated. Recently, Miller¹ presented arguments that the particles and pits represent conical deformations of membranes, rather than inverted micelles sandwiched between membrane leaflets—a model repeatedly suggested by Verkleij *et al.*²⁻⁴. We have also observed lipidic particles and pits in liposomes of mixed egg phosphatidylcholine (PC)–soybean phosphatidylethanolamine (PE)⁵ and mixed PC–PE–cardiolipin, and our results essentially agreed with those of others¹⁻⁴. Moreover, we have more convincing evidence for the deformation model as presented below. The evidence has led us independently to arrive at a similar model⁶ to that proposed by Miller.

We have obtained cross-fractured profiles of the particles and pits in multilamellar vesicles. As shown in Fig. 1a, most of the remaining halves of the fractured particles and pit are seen as interlamellar contact points, although a few single particles (open arrows) are also observed. The profile of the lipidic particle is clearly conical, with the base extending to 200 nm. They are considerably larger than the mean size of proteinaceous intramembranous particles, as seen in Fig. 1b. Here, in a PC–PE vesicle containing human erythrocyte membrane protein, large conical lipidic particles are clearly seen at the fracture line as contact sites between adjacent membranes. Proteinaceous particles are hemispherical with diameters of ~10 nm. As the lipidic particles and pit are really cusp-shape distortions, they should be named as such instead. When multilamellar vesicles containing <30% PC in PE were quenched at room temperature, rows of cusps on the fractured faces were frequently observed. However, cusps are not found in large single lamellar vesicles (verified by negative staining) of the same lipid mixture. This provides independent evidence that intermembrane attachment is required for the formation, stabilization and alignment of the cusps. Our results give definite support for the cusp model¹ in preference to the inverted micelle model².

This work was supported by NIH grant GM 28120 and American Cancer Society grant BC-248.

S. W. HUI
THOMAS P. STEWART

Electron Optics Laboratory,
Biophysics Department,
Roswell Park Memorial Institute,
Buffalo, New York 14263, USA

1. Miller, R. G. *Nature* **287**, 166–167 (1980).
2. de Kruijff, B. *et al.* *Biochim. biophys. Acta* **555**, 200–209 (1979).
3. Verkleij, A. J., Mombers, C., Gerritsen, W. J., Leunissen-Bijvelt, L. & Cullis, P. R. *Biochim. biophys. Acta* **555**, 358–361 (1979).
4. Verkleij, A. J., van Echeld, C. J. A., Gerritsen, W. J., Cullis, P. R. & de Kruijff, B. *Biochim. biophys. Acta* **600**, 620–624 (1980).
5. Hui, S. W., Stewart, T. P., Yeagle, P. C. & Albert, A. *Archs Biochem. Biophys.* (in the press).
6. Hui, S. W., Stewart, T. P., Boni, L. T. & Yeagle, P. L. *Science* (in the press).

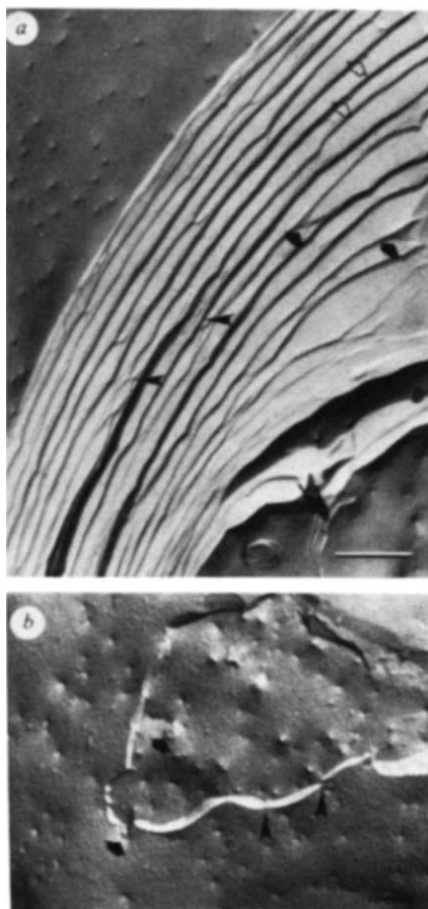


Fig. 1 Freeze-fracture electron micrographs of multilamellar vesicles of 25% PC in PE (a), and with added crude human erythrocyte membrane protein (smaller particles) extracted with Triton X-100 (b). Lipidic particles and pits are seen as intramembrane attachment sites in cross-fractured areas (arrowhead), some of which have developed extensive contacts (solid arrows). Open arrows indicate occasional conical particles on one bilayer alone. Experimental procedures are given elsewhere^{5,6}. Scale bar, 500 nm.

VERKLEIJ AND DE KRUIJFF REPLY—We believe that 'lipidic particles' represent inverted micelles formed either within a single bilayer^{1,2} or at the nexus of intersecting bilayers during fusion of membranes³⁻⁵. We have presented^{2,3} evidence for this interpretation and

against the lipidic particles being local bilayer curvatures as inherent to the intermembrane attachment side (IMAS) model as proposed by Miller⁶ or the similar cusp model proposed by Hui and Stewart in their letter. These arguments and new facts supporting our hypothesis can be summarized as follows:

(1) The crucial point ignored by the IMAS model is that the lipidic particles have been exclusively observed in lipid systems in which at least one of the lipid species prefers the hexagonal phase II (H_{II}). In this latter phase the lipid molecules are cone shaped and oriented with their polar head group towards an aqueous channel of a small radius. Any model of the lipidic particle must take into account this favourable organization of H_{II} -type lipids, a criterion met in the inverted micelle model. This requirement, however, is not fulfilled with the IMAS model. For instance, in the cardiolipin–lecithin system⁶, cardiolipin will be present at the top of the strongly curved bilayer where it can only fit if the molecules have an 'inverted cone' shape. However, cardiolipin assumes a cone shape in the presence of Ca^{2+} .

(2) In line with this argument, a direct morphological association has been observed between the H_{II} phase and the lipidic particles^{4,7}, indicating a close relationship of these structures. Furthermore, we have found in some systems^{3,7} three-dimensional arrays of lipidic particles which probably represent an agglomerate of inverted micelles. An analogous example of structural relationships between micellar and hexagonal phase can be found in another area of lipid polymorphism in that lipids can adopt both micellar and hexagonal H_I phases in water, as, for instance, does lysophosphatidylcholine⁸.

(3) In the IMAS model, the bilayer must be strongly curved as the lipidic particles can be very small (60–130 Å in diameter)². Such a strong bilayer curvature is energetically highly unfavourable because the minimum curvature of the bilayer in the most strongly curved system (sonicated vesicles) is at least 250 Å in diameter.

(4) The IMAS model ignores the extensive ³¹P NMR studies on the lipidic particle-containing systems^{2,9}. The profound hysteresis of the isotropic peak in the ³¹P-NMR spectrum of these systems can be best explained by a model in which inverted micelles are located in rows at the fusion site of two bilayers¹⁰.

In conclusion, we feel that these arguments strongly suggest that lipidic particles represent intrabilayer inverted micelles. Also, the observation that in systems such as that described by Hui and Stewart the particles are relatively large and cone shaped can be interpreted as the

presence of inverted micelles at the fusion site of two well separated bilayers¹¹.

A. J. VERKLEIJ
B. DE KRUIJFF

*Institute of Molecular Biology,
State University of Utrecht,
PO Box 80.063,
3508 TB Utrecht,
The Netherlands*

1. Verkleij, A. J., Mombers, C., Leunissen-Bijvelt, J. & Ververgaert, P. H. J. Th. *Nature* **279**, 162–163 (1979).
2. de Kruijff, B. et al. *Biochim. biophys. Acta* **555**, 200–209 (1979).
3. Verkleij, A. J., Mombers, C., Gerritsen, W. J., Leunissen-Bijvelt, J. & Cullis, P. R. *Biochim. biophys. Acta* **555**, 358–362 (1979).
4. Verkleij, A. J., van Echteld, C. J. A., Gerritsen, W. J., Cullis, P. R. & de Kruijff, B. *Biochim. biophys. Acta* **600**, 620–624 (1980).
5. de Kruijff, B., Cullis, P. R. & Verkleij, A. J. *Trends biochem. Sci.* **5**, 79–81 (1980).
6. Miller, R. G. *Nature* **287**, 166–167 (1980).
7. van Venetie, R., Leunissen-Bijvelt, J. & Verkleij, A. J. *Biochim. biophys. Acta* (submitted).
8. Luzzati, V., Reiss-Husson, F., Rivas, E. & Gulik-Krzywicki, T. *Ann. N.Y. Acad. Sci.* **137**, 409–413 (1966).
9. Cullis, P. R. & de Kruijff, B. *Biochim. biophys. Acta* **559**, 399–420 (1979).
10. Cullis, P. R., de Kruijff, B., Hope, M. J., Nayar, R. & Schmidt, S. *Can. J. Biochem.* **58**, 1091–1100 (1980).
11. Verkleij, A. J. et al. (in preparation).

Fossil charcoal and the palaeoatmosphere

It has been proposed by Cope and Chaloner¹ that oxygen levels in palaeoatmospheres may be assessed by the presence or absence of charcoal in the geological record. Although oxygen is required for burning, it is not necessary for charcoal formation. When wood is subjected to high temperatures, oxygen-bearing gases are evolved, yielding an oxygen-depleted graphitic network (charcoal)². Charcoal formation is favoured by low oxygen concentrations at the solid surface and forms in anoxic conditions at temperatures above 200 °C. In the remote past, however, heat from wood fires, sustained by atmospheric oxygen, was probably the dominant agent for charcoal formation. Thus, the coupling of atmospheric oxygen concentration to charcoal formation may not be as direct as has been proposed.

That the occurrence of charcoal is determined by the oxygen concentration required for carbon monoxide or methane to burn, as Cope and Chaloner presume, has not been demonstrated. Depending on configuration, ignited wood samples may not continue to burn in atmospheres containing as much as 13–21% oxygen³. At lower concentrations the most likely source of heat for charcoal formation would not be available.

It is not known whether palaeoatmospheric pressures were identical to present-day values. The proportion of oxygen required for ignition depends on the pressure⁴, but the dependence has not been investigated for wood. If inert gas is added to air, as in submarines, the partial pressure of oxygen and consequently

respiration are unaffected, but the proportion of oxygen may be reduced sufficiently to prevent a match from burning⁵.

While applauding the cross-disciplinary approach of Cope and Chaloner, we caution that the effects we have described should be considered before this potentially valuable probe of palaeoatmospheric composition be used. With refinement, this might be extended to explore fascinating problems, such as atmospheric implications of expanding Earth models and the effects of oxygen and carbon dioxide partial pressures on photosynthesis^{6–8} in the past.

F. R. S. CLARK

*Fire Research Section,
National Research Council
of Canada,
Ottawa, Canada K1A 0R6*

D. A. RUSSELL

*Paleobiology Division,
National Museum of Natural
Sciences,
Ottawa, Canada K1A 0M8*

1. Cope, M. J. & Chaloner, W. G. *Nature* **283**, 647–649 (1980).
2. Browne, F. L. *Report 2136* 18–19 (Forest Products Laboratory, Madison, Wisconsin, 1958).
3. Rasbash, D. J. & Langford, B. *Combustion Flame* **12**, 33–40 (1968).
4. Wharton, R. K. *Fire Mater.* **3**, 39–48 (1979).
5. Carhart, H. W. *Nat. Symp. Fire Aspects of Polymeric Materials* (American Chemical Society, June, 1977).
6. Gardescu, J. & Hagar, W. G. *Pl. Physiol.* **65**, Suppl. 6, 12 (1980).
7. Vidaver, W. & Hall, G. *Pl. Physiol.* **65**, Suppl. 6, 13 (1980).
8. Rogers, H. H., Sinclair, T. R. & Heck, W. W. *Pl. Physiol.* **65**, Suppl. 6, 49 (1980).

CHALONER AND COPE REPLY— Although we are aware that charcoal may be produced in a variety of physical conditions, we must stress that the theme of our letter was to examine the palaeoatmospheric implications of our studies of the widespread distribution of charcoal (fusinite) in the geological record. We did not attempt to present a comprehensive survey of the conditions of charcoal formation.

In the natural environment, charcoal may be produced in conditions of both pyrolysis and combustion, but the mechanisms for charcoal formation are few. First, basalt lava flows are known to engulf vegetation and subject the included woody tissues to pyrolysis, transforming these into charcoal¹. Second, in swamp peat accumulations, local temperature rise caused by exothermic decomposition reactions may initiate smouldering which could produce charcoal. Third, and most important, combustion of vegetation by forest fire can result in the production of large volumes of charcoal as a residue. It is the last mechanism which we believe to be responsible for the generation of the greater proportion of charcoallified phytoclasts which occur in ancient sediments. Thus, our suggested implication for past atmospheric composition is not based on the range of physical conditions for charcoal production but on the

restriction (in terms of oxygen level) which can be placed on the occurrence of large-scale natural burning, that is, the combustion associated with forest fire. We placed this restriction at a minimum oxygen level of ~0.3 present atmospheric level (PAL). We acknowledge that it may be possible to produce charcoal in special circumstances with oxygen levels <0.3 PAL. However, at such levels it would not be possible to maintain the combustion mechanism for the production of the large quantity of charcoal preserved in the geological record, estimated to be 4×10^5 – 4×10^7 tons in post-Silurian rocks².

Since the publication of our letter we have learned (K. N. Palmer, personal communication) that the combustion levels of carbon monoxide and methane described in the experiments of Coward and Jones³ were based on pre-mixed flames. In conditions of natural burning the combustion is by diffusion flames, where the minimum level of oxygen required to support combustion is more likely to be ~13% (~0.6 PAL), corresponding to the lower levels of oxygen concentration in the extinction experiments of Rasbash and Langford⁴. Thus, we are able to revise our minimum estimate to ~0.6 PAL, thereby placing further constraints on estimates of oxygen concentration in post-Silurian atmospheres.

The question of increased atmospheric pressures is interesting. Since it is generally accepted that the oxygen in the Earth's atmosphere has accumulated steadily from photosynthetic build-up, any proposed models of palaeoatmospheres should set the partial pressure of oxygen at or about its present level. Presumably, for an oxygen/nitrogen palaeoatmosphere at higher than present-day atmospheric pressure, the implied increased partial pressure of nitrogen would have to be accounted for by increased rates of volcanic outgassing in the past. We do not know of any geological evidence to support or reject this type of model, and therefore can make no further comment on this aspect. In the absence of such data we can only assume that past atmospheric pressures were similar to those of the present day. In such conditions we offer the occurrence of charcoal in the sedimentary record as a valuable geological indicator in any proposed palaeoatmospheric models.

M. J. COPE

W. G. CHALONER

*Department of Botany,
Bedford College,
Regent's Park,
London NW1 4NS, UK*

1. Lockwood, J. P. & Williams, I. S. *Geol. Mag.* **115**, 69–74 (1978).
2. Cope, M. J. in *Organic Maturation of Sedimentary Organic Matters. Preservation Problems and Fossil Fuel Exploration* (ed. Brooks, J.) (Academic, New York, in the press).
3. Coward, H. F. & Jones, G. W. *Bull. U.S. Bur. Mines* **503** (1952).
4. Rasbash, D. J. & Langford, B. *Combustion Flame* **12**, 33–40 (1968).

BOOK REVIEWS

Doing things in the public interest

Eric Ashby

IT TAKES courage to write a book called *Implementing Public Policy*. The author walks in the footsteps of some very distinguished stylists: Bentham and Bagehot in the last century, Wheare and Aron in this century. It was Bagehot who made the point that success in public policy "depends on a due mixture of special and non-special minds — of minds which attend to the means and of minds which attend to the end". Since Bagehot's time the means have become more sophisticated and the end more controversial. The "due mixture" of special and non-special minds sometimes fails to produce the desired result. As a rule the special mind produces a rational case, based on hard data; the non-special mind employs what is called political judgement, which is just a dignified way of saying "hunch". Lord Salisbury, writing over a hundred years ago, made the point so well that there is no room for improvement: "... you should never trust experts. . . . They all require to have their strong wine diluted by a very large admixture of insipid common sense".

Ruth Levitt's book is a variation on this theme. The merit of her book is that she eschews vague generalization about how a decision to do something in the public interest becomes transformed into something actually done in the public interest. Instead she selects a specific example of public policy and uses this as a paradigm for public policy in general. Her specific example is policy for pollution. It is an apt choice: all political parties agree that it is a good thing to abate pollution, so the issue is not bedevilled by party politics. The means for abating pollution are highly technical and cannot be left to politicians or bureaucrats; the "due mixture" has to include scientists, economists and lawyers. And the end, though not entirely free from controversy, is not obscured by ideologies. The choice of pollution, therefore, clears the ground of several difficulties.

Ruth Levitt's technique is to dissect the *Control of Pollution Act, 1974*, and then to examine its implications in four case-studies: toxic waste disposal, the EEC directive on bathing water, noise abatement, and industrial air pollution. In fact — and Ruth Levitt makes this clear — these case-studies involve not only the *Control of Pollution Act* but other laws as well. She describes, with abundant references to recent literature, how the control of dumping of toxic wastes arose in response to sensational reports in the mass media, not in response to the considered recommend-

Implementing Public Policy. By Ruth Levitt. Pp.213. (Croom Helm/Biblio, Totowa, NJ: 1980.) £11.95, \$30.

ations of a committee appointed to examine the problem by the Ministry of Housing and Local Government. The advice of the special mind was disregarded. It was "insipid common sense" on the part of the non-special mind which, in 1972, hustled a Bill through Parliament in record time to control poisonous waste.

Section 17 of the *Control of Pollution Act* is intended to supersede this hasty legislation; but action on this was delayed for over six years, partly because the special minds could not define precisely what toxic waste was. At long last, on 17 November 1980, regulations to control special waste were laid before Parliament. They illustrate one of the themes of Ruth Levitt's book: the difficulty of getting unambiguous and politically viable advice from the special minds. The regulations proscribe the uncontrolled dumping of wastes "dangerous to life". And how do you think "dangerous to life" is defined? Waste is dangerous to life if "a single dose of not more than five cubic centimetres would be likely to cause death . . . if ingested by a child of 20 kilograms' body weight". It is no wonder that there has been a prayer against these regulations in the House of Lords, where the criterion of dangerous-to-life was ridiculed in a spate of not-so-insipid common sense (one noble baroness called it "cock-eyed").

Ruth Levitt's second case-study is another example of the mismatch between the specialist and generalist ingredients of a political decision. The EEC's policy is to harmonize environmental laws in the member nations; so the Commission issues inflexible directives which are intended to apply from the tip of Scotland to the tip of Sicily (and now to the isles of Greece). The directive on bathing water, obviously drawn up with an eye to a crowded resort on the tideless Mediterranean, prescribes mandatory values for some dozen parameters that have to be measured regularly, especially where "the concentration of bathers exceeds a mean value of 10,000 persons per linear kilometer". If the state of British bathing beaches were a hazard to health, such a draconian system of monitoring might be justified; but there is no evidence whatever to support the view that the very great cost of monitoring bathing beaches in Britain according to this

directive is justified. Yet it is now Community Law; an example of policy-making just for the sake of policy-making.

The other two case-studies are equally illuminating. Policy on noise abatement, like policy on the deposition of toxic waste, is hampered by insufficient technical knowledge. Policy on industrial air pollution is not hampered in this way but it is misunderstood. The achievements of the Alkali and Clean Air Inspectorate are in fact one of the great success stories in the control of pollution. But there has never been adequate communication between the Inspectorate and the public. There is provision in the *Control of Pollution Act* (sections 79–83) to empower local authorities to keep registers, open to public inspection, with particulars of industrial emissions. A Royal Commission, a committee set up by the Clean Air Council, and sundry groups of environmentalists have urged that the public have a right to know what is in the air they breathe, but practically nothing has been done; an example of another point Ruth Levitt makes, that there is a gap — sometimes a deliberately unbridged gap — between the passing of a law and its implementation.

The case-studies are valuable and, on the whole, accurate. (Some of Ruth Levitt's statements are not, in my view, correct, but it would be quibbling to dispute them.) What is disappointing about the book is its neglect of the historical dimension of pollution policy, and the omission of any reference to studies on similar public policy done in the USA and Canada. The purpose of the case-studies is to illustrate the general pattern of decision-making for public policy. But this pattern cannot be understood if it is observed without reference to its evolution. Thus the present interpretation of the formula "best practicable means", discussed in Chapter 7 of the book, makes little sense if one does not take account of the early use of the formula in the Towns Improvement Acts in the early 1800s, the criticism of these by de la Beche and Playfair in their pioneer report to the government in 1846, and the new interpretation given to it by the Alkali Inspectorate in the 1870s.

As for reference to work on the other side of the Atlantic, I think Ruth Levitt's concluding chapter on effective public policy would have been greatly improved if she had taken account of the work done by Dorothy Nelkin and her associates in Cornell. Ruth Levitt mentions the role of

public opinion in agenda-setting, but she barely mentions the implications so perceptively analysed in *Technological Decisions and Democracy* (see *Nature* 280, 849; 1979) and other papers by Nelkin.

There is still great confusion about how public policy is made. The present demand for public participation in policy about the environment makes it very important that this confusion should, as far as possible, be cleared away. Ruth Levitt's book is a

serious and useful contribution toward this end. But, as the author herself says in the preface, her idea has not been to make judgements, but rather "to observe, clarify and draw attention to relevant detail". This is what she has done. □

Lord Ashby is Chancellor of Queen's University, Belfast, and a Fellow of Clare College, Cambridge. He was a member of the Royal Commission on Environmental Pollution from 1970 to 1973.

Molecular genetics: a language to be learnt

Jirí Novotný

Evolution and Variation of Multigene Families. Lecture Notes in Biomathematics, Vol. 37. By Tomoko Ohta. Pp. 131. (Springer-Verlag: 1980.) Pbk DM 19.50, \$11.50.

GALILEO Galilei said that mathematics is the language in which God wrote the book of nature. Most of science seems to substantiate this idea, but there are exceptions. The whole field of chromosomal DNA study, for example, presents us with a conspicuously non-quantitative picture. Although this feature may contribute to its present-day popularity, it is also rather disturbing. It signals that our contemporary knowledge of gene and chromosome organization (to say nothing of chromosome evolution) is far from a real understanding and is merely superficial at best.

Only partly is this situation due to a lack of quantitative information. Since the advent of DNA cloning and sequencing, volumes of such data are emerging almost daily. What we seem to be lacking, however, is quantitative *understanding*. Quantitative data are no more useful than qualitative if we do not know what they mean and in what context to understand them. The uniqueness of Ohta's book lies in the fact that, for the first time, it outlines a quantitative theory of what happens with the genomic DNA in the course of time.

The book is brief, informal and although it contains some (rather unimportant) errors, it is appealing in its sincere endeavour to provide us with a unified and thorough explanation of diverse phenomena: genetic correlation among and within multigene families governed by unequal crossing-over, amino acid sequence variation of immunoglobulins, amino acid site linkage disequilibrium, gene conversion, accumulation of non-functional genes. This is a difficult goal; in consequence the book is sometimes hard to follow. Compared to the intricacy of the subject itself, it is nevertheless a condensed and understandable presentation which is well worth reading. Ohta's work goes far beyond the pioneering studies of Smith

(*Science* 191, 528-535; 1976) and others in that it deals not only with the evolution of a single chromosome but also with the sexual gene exchange in a population. A key notion of the theory is the identity coefficient: the probability of two genes of the same multigene family being identical by descent. Several different coefficients are defined and compared to analyse different levels of the genome; conclusions drawn from their comparison are indeed interesting. Particularly original and innovative is the analysis of immunoglobulin V regions. The within-species and between-species coefficients are computed separately for the framework and hypervariable sections of the V genes and shown to vary correlatively. Another part of the book gives a non-traditional discussion of a notorious issue, the question of V gene families and subfamilies and their interrelation in the course of evolution.

I have no doubt that anybody who is interested in molecular evolution will find this book stimulating and useful. □

Jirí Novotný is an Assistant Professor in the Cellular and Molecular Research Laboratory of Massachusetts General Hospital and Harvard Medical School.

Great balls of fire?

E. Philip Krider

Ball Lightning and Bead Lightning: Extreme Forms of Atmospheric Electricity. By James Dale Barry. Pp. 298. (Plenum: 1980.) \$29.50, £18.59.

THE remarkable phenomenon of ball lightning has both frightened and fascinated mankind for centuries. A typical ball lightning is a mobile sphere, about the size of an orange or grapefruit, which is produced during a thunderstorm and which remains luminous for one or more seconds. *Ball Lightning and Bead Lightning* is the second book to be published about this subject by Plenum in the past ten years. The first book, *The*

Nature of Ball Lightning by S. Singer, was an attempt to integrate many of the previous papers on ball lightning, both observations and theories, that have appeared over the past several hundred years.

This second work by J. D. Barry reviews the known physical characteristics of ball lightning that have been deduced from visual observations and laboratory experiments, and then attempts to deduce other characteristics and properties which are consistent with these observations. One chapter is also devoted to a discussion of bead lightning, a more familiar phenomenon but one which is still poorly documented in the scientific literature.

The book begins with a brief description of normal lightning and an assumption that ball and bead lightning do indeed exist. Chapter 2 lists all known references to bead lightning, and then presents all published photographs of this phenomenon together with a critical evaluation of their interpretations. The remaining six chapters are devoted to a review of the characteristics of ball lightning, a presentation and critical evaluation of all published photographs that have been reported to be of ball lightning, a review of the laboratory experiments which have attempted to reproduce the natural event and a comprehensive bibliography.

The book is well written and complete at a qualitative level. The principal deficiencies are in the review of normal lightning characteristics and the discussion of bead lightning — the characteristics of normal discharges are quite important if one is to evaluate production mechanisms or if one wishes to compare ball characteristics with a normal channel. No values are given for leader and return stroke currents or their statistics; there is no discussion of continuing currents; and the electric power and energy dissipated by normal flashes is not considered. There is no presentation of time-resolved photographs, which almost always show the streaks of persisting luminosity that may well be a short-duration form of bead lightning. The suggestion that bead effects are most commonly observed during cloud-to-cloud discharges (p.11) is contrary to the cloud-to-ground emphasis in works previously published by lightning researchers. There is also no discussion of whether a simple projection of the three-dimensional channel tortuosity toward or away from the observer could explain the beaded appearance of, say, a long-continuing current.

Despite these shortcomings, the book is a valuable introduction to a fascinating subject and will certainly stimulate research in this area. The collection of over 50 photographs and a bibliography of over 1600 entries are truly exceptional and will serve the scientific community well for years to come. □

E. Philip Krider is a Professor at the Institute of Atmospheric Physics, University of Arizona, Tucson.

Cooperation in limnology, a further achievement of the IBP

C.M. Yonge

The Functioning of Freshwater Ecosystems. Edited by E.D. Le Cren and R.H. Lowe-McConnell. Pp.588. (Cambridge University Press: 1980.) £40, \$95.

THE objective of the International Biological Programme was the furtherance of our understanding of the biological basis of productivity and its relation to human welfare. The study of productivity involved the organization of international studies on land, in the sea and in freshwaters. The last-named posed the greatest problems, both because previous international studies had been slight and because conditions in such waters range so very widely — from oligotrophic highland lochs to tropical alkaline lakes containing dense but specifically restricted algal populations which support equally specialized fish and bird populations.

Because it involved the stimulation of self-financing local enquiries, no section of the programme demanded greater organization. The eventual world-wide participation is indicated in the maps which adorn the end pages of this book which, following earlier publications concerned largely with methodology, represents the final outcome of the freshwater programme or, more concisely, IBP/PF. Under the overall supervision of the Director General, Dr Barton Worthington, especial responsibility rested with the full-time coordinator, Dr Julian Rzoska, to whose dedicated activities, including the particularly close involvement of eastern European countries, this impressive volume is a tribute.

As Professor Livia Tonelli, the last of a number of convenors, writes in the Introduction, "This book is the result of years of a corporate effort by many scientists all over the world to advance the knowledge of production limnology". The truth of this statement is borne out when it is realized that it represents the combined contributions of no less than 43 authors reporting from the 93 widely scattered sites. Not only was interest so widely stimulated but important bilateral projects came into being between such widely separated countries as West Germany and Brazil, France and Chad, Japan and Malaya, and Britain and Uganda.

Freshwaters vary widely in character and consequent productivity; they are also quickly and often profoundly influenced by local human activities with consequences ranging from total toxicity to extreme eutrophication. Yet they also have great potential, realized most fully in Southeast Asia, for the production of fish and prawns for human consumption.

All means of control, alike against damage and towards increased productivity, depend on basic understanding of prevailing conditions. It is this that is the

theme of the book. The chapters range from discussion of the effects of physical and chemical variables to analysis of primary and secondary production. Studies on the significance of organic matter and decomposers are followed by chapters on trophic relationships and efficiencies, then come others dealing with estimations of productivity and consideration of the value of dynamic models of lake productivity. The final

chapter by G.G. Winberg provides a most interesting summary of the results of Soviet freshwater studies.

The overall result with its final 66 pages of references is a volume of unique value to all limnologists and one of the major achievements of the International Biological Programme. □

Sir Maurice Yonge is an Honorary Fellow in Zoology at the University of Edinburgh.

Wooldridge and Linton re-viewed

F. W. Shotton

The Shaping of Southern England. Institute of British Geographers Special Publication, No 11. Edited by David K. C. Jones. Pp.274. (Academic: 1980.) £14.60, \$34.50.

THIS book was commissioned by the Institute of British Geographers to celebrate the fortieth anniversary of the publication of Wooldridge and Linton's *Structure, Surface and Drainage in South-East England* (IBG, 1939). In a preface and ten chapters, eleven authors are involved, some of them more than once. This assembling of the latest opinions of geomorphologists on the evolution of the topography of their own terrain, namely the Thames Basin, the Weald and the Hampshire Basin, will be appreciated by geologists and geographers from a much wider area. The evolution of the form and drainage of this part of England was argued about long before Wooldridge and Linton, as K. Clayton points out. The value of *Structure, Surface and Drainage* lay less in its conclusions than in the way it combined meticulous field observation with scientific interpretation and thus steered geomorphological thought to its present position.

Although it is stated that the book is a memorial to Wooldridge and Linton, only four chapters examine their conclusions critically. Two of these, by D. K. C. Jones and R. J. Small respectively, deal with the development of the landscape during the Cainozoic era, and although Small adds the words "an alternative interpretation" to his title, it is more of an alternative to Wooldridge and Linton than to Jones. Both authors reject two of the main tenets of the original interpretation, which postulated a late Oligocene or early-Miocene "Alpine storm" of earth deformation with subsequent tectonic quiescence while the land surface was reduced to a peneplane, and a marine invasion (the so called Calabrian or late-Neogene transgression) to a height of about 200m. Instead they envisage a

succession of deforming pulses from the early Eocene onwards, with concurrent erosion or deposition.

Two other papers, by C. A. Baker with D. K. C. Jones and by C. P. Green and D. F. M. McGregor, deal with the diversions of the course of the Thames during the Pleistocene. These I found less satisfying, for although the authors differ from several points of the original interpretation, they also reveal how much still remains unresolved. On the age and origin of the Chiltern Drift, surely pre-Cromerian but not inevitably a till, there is no clear consensus.

The other contributions are not so closely related to Wooldridge and Linton but are no less interesting because of that. They deal with the origin of the sarsens, the residual deposits of the North Downs, the palaeohydrology of the Kennet Valley and the effect of periglacial weathering on Chalk. A final chapter by C.P. Green describes the development of geomorphology as it is now understood and practised, and discusses its probable future. Green points out that modern geomorphology, combining as it does the study of land form and drainage, the physical processes by which these are developed, stratigraphy, sedimentology, palaeoenvironment reconstructions and relative and absolute chronology, was taken over by geographers from the geologists. It gives no pleasure to me to read (and half admit) that this was due to "the relative lack of interest in the Quaternary among British geologists". In fact Wooldridge, acknowledged as one of the great pioneers in modern geomorphology, was, by training and early employment, a geologist. Green's chapter emphasizes that geomorphology is now neither simple geography nor geology but a hybrid of both, biology also providing a formative influence. The subject is truly scientific and very much a part of earth science. □

F. W. Shotton is Emeritus Professor of Geology at the University of Birmingham.

CORRESPONDENCE

Continued from page 356

of confusing the origins with the content of a scientific theory⁴. Any student of sociology or philosophy of science is likely to be intrigued by the social framework of the emergence of sociobiological theory over the last decade and its relationship to the New Right ideology⁵. But having explored that link, it is the task of critics of sociobiology to expose what they see as its methodological, internal inadequacies (for example, refs 6 and 7).

If sociobiologists want to avoid the charge that they believe that biology is destiny, they should beware of telling magazines that they know "why we do what we do" or entitling their books *The Selfish Gene*, *The Inevitability of Patriarchy* or *On Human Nature*. The trouble is that they want to have their cake and eat it. They imperialize the human sciences (*vide* the first paragraph of *Sociobiology*, *The New Synthesis*) and are embarrassed by the outcome. This is why the final chapters of Dawkins' or Wilson's books are so confusing. Having set out the inexorable destiny of genetic predispositions to xenophobia, aggression, patriarchy or whatever, they invoke the possibility of a human conscious prospect of overcoming these predispositions. As Dawkins puts it in his letter, with respect to philandering males "many humans (have) some at least momentary intention of overcoming their polygamous tendencies. Many even succeed in this". For Wilson eugenics may overcome any "hereditary tendency" to acquire xenophobia. Free will, intentions and wishes (or Dawkins' memes) like the US cavalry, come galloping over the horizon in the nick of time to rescue us from our genes.

But *where* does our free will, etcetera come from? How can we be *both* genetically programmed robotic DNA survival machines and have the extraordinary capacity to transcend these programs? The truth is that sociobiological determinism, when challenged, collapses into the weakest sort of Cartesian dualism. For consistent materialists — like Gould, whom Dawkins quotes, and, I hope, myself too — we must argue that wishes, intentions etc. are as much, or as little, "given" by our genes as any other aspect of our human existence. The interesting thing about humans is that the fact that we can change what we do is as much part of our biology as how we do what we do. To give the example used in my review of Wilson — if humans were quadruped and not biped, their social arrangements would be different; that humans are biped is genetically "given", therefore the human social order is genetically given. The point is that such genetic syllogisms are boringly uninteresting about either any of the crucial aspects of differences *between* human societies or the changes that occur *within* any given society. They neither explain nor predict apartheid in South Africa, cultural revolutions in China, Born Again Christianity in the United States, the welfare state or its dismantling under Mrs Thatcher in Britain, still less any of our individual proclivities.

On such questions then, biology has to be silent. It is because self-styled sociobiologists are *not* silent (I challenge anyone to read *The Selfish Gene* and come away *without* a clear impression of Dawkins' views of what biology

has to say about the Welfare State, sexual mores or microeconomics) that their work both trivializes important social and biological questions and is so amenable to neo-Nazi and New Right ideology.

STEVEN ROSE

*The Open University
Milton Keynes, UK*

1. Maynard Smith, J. *Nature* **289**, 742 (1981).
2. Dawkins, R. *Nature* **289**, 528 (1981).
3. Gould, S.J. *Nature* **289**, 742 (1981).
4. Rose, H. & Rose, S. *Commun. Cognition* **13**, 269-283 (1980).
5. Rose, H. & Rose, S. in *Markets and Welfare* (eds Burkitt, B. & Davey, A.) (in the press).
6. Gould, S.J. & Lewontin, R.C. *Proc. R. Soc. B* **205**, 581-598 (1979).
7. Montagu, A. *Sociobiology Examined* (Oxford, 1980).

Conservation sites

SIR — The attack on conservation sites by Muir (*Nature* 12 March, p.82, 173), apart from ignoring that the scientific value is by no means the only value considered by conservationists, is based on several unacceptable assumptions.

In implying that conservation sites are primarily for the preservation of rare species Muir ignores the fact that most have been primarily designated in order to conserve particular vegetation types, associations of organisms, natural or semi-natural ecosystems. The argument that "value to science of any objects or phenomena lies not in themselves but in the information they yield to study. Once this information is recorded and published whatever value remains in the objects or phenomena is of no value to science" is ridiculous on two counts.

First, he specifies no criteria of "value" that are meaningful. He merely transfers the problem of what he means from conservation sites to published information about them. Even if one disregards this problem one is left, on his own arguments, with a strong case in favour of conservation sites. Ecology is still a young science and it has hardly scratched the surface in its study of the vast majority of ecosystem types. The rate of ecological advance does not compare favourably with the rate of loss of certain habitat types. There are, as yet, no vegetation types or ecosystems whose study has been exhausted to the extent that they are of no further interest to ecologists!

I am responsible for the "sites of special scientific interest" that make up the Malham Tarn Nature Reserve, one of the best documented nature reserves in the country. Its scientific documentation started with observations by John Ray in 1671 and has rapidly accelerated since 1947. We now know enough to realize that it will require much more study than that already accomplished before any scientist would be other than foolish to suggest the site was of no further interest to science.

The tragedy of the present rate of loss of sites of special scientific interest is that we know so little of what is being lost. What little we do know strongly suggests that the Nature Conservancy Council have correctly identified the sites of greatest interest to science apart from them being those necessary if we are to achieve the conservationists' aim to convey the maximum diversity of wildlife into the next century. The conservationist is not, as Muir implies, trying to halt evolution. He is

primarily concerned with the perpetuation of a diversity of habitat types. A variety of forest organisms depends on a variety of forest habitats — regardless of the rates of evolution of their organisms.

R.H.L. DISNEY

*Malham Tarn Field Centre,
Settle, North Yorkshire, UK*

DMSO and immunity

SIR — The observations of Pestronk and Drachman¹ that dimethyl sulphoxide (DMSO) reduced anti-receptor antibody titres in experimental myasthenia gravis in rats are of considerable interest. The authors conclude: "It seems likely that DMSO might also be effective in ameliorating other active humoral immune responses. If so, DMSO may provide an effective new treatment for immune diseases mediated by autoantibodies." Lest this statement lead once more to proclamations of DMSO as a wonder drug, several important points need an answer.

The LD₅₀ of DMSO by intraperitoneal (i.p.) injection is between 10 and 11 per kg for rats and mice. Pestronk and Drachman injected 1.0 ml of DMSO into rats daily for 14 days. Assuming a body weight of 250 g for a rat the dose of DMSO for each rat would be approximately 4 g per kg, or 36 per cent of the acute LD₅₀. The injection (i.p.) of this amount of DMSO into rats and mice produces a number of pharmacological and biological effects including an increase in peripheral blood pressure, changes in the oxygen levels, probable anoxia in the spleen and hypothermia^{2,3}. In mice the level of hypothermia is severe, with rectal temperature reaching 33.5°C and recovery after a single injection of DMSO (4.5 g per kg) takes more than 6 h. Hypothermia represents a generalized but reversible toxicity. Many chemicals when injected at levels of approximately 20–40 per cent of the LD₅₀ produce similar effects.

Thus the clearly demonstrated effects of DMSO in decreasing antibody titres illustrated by Pestronk and Drachman may not be specific. In any event the possible therapeutic usefulness of a chemical effective at doses so close to the LD₅₀ is unlikely to prove clinically acceptable.

M. J. ASHWOOD-SMITH,

*University of Victoria,
British Columbia, Canada*

1. Pestronk, A. & Drachman, D. B. *Nature* **288**, 733-734 (1980).
2. Ashwood-Smith, M. J. Thesis, Univ. London (1962).
3. Ashwood-Smith, M. J. *Int. J. Radiat. Biol.* **3**, 101 (1961); *Ann N. Y. Acad. Sci.* **141**, 45–62 (1967).

SIR — The comments of Professor Ashwood-Smith should serve as a useful reminder that the pharmacological actions of dimethyl sulphoxide must be carefully examined. The mechanism of action by which the drug suppresses anti-acetylcholine receptor antibodies in the experimental animal model of myasthenia gravis is not yet understood. Moreover, the reduction in serum antibody levels must be taken into account as a possible undesirable effect if the drug is contemplated for other clinical applications.

ALAN PESTRONK

DANIEL B. DRACHMAN

*School of Medicine,
John Hopkins University,
Baltimore, Maryland, USA.*